1 The Model

An *applicant* (he) with either *High* or *Low* quality, $\theta \in \{H, L\}$, seeks an approval from any one of n *evaluators* (she). To obtain the approval he seeks, he sequentially visits these n evaluators, each at most once. Any evaluator who receives his visit decides whether to *approve* or *reject* the applicant. Once he is approved or has visited all n evaluators, the applicant stops his visits. Otherwise he visits the evaluator labelled $\tau(k)$ after his $k-1^{\text{th}}$ rejection, where $\tau(.)$ is a permutation of the set of evaluators' labels $\{1, 2, ..., n\}$.

If an evaluator approves the applicant, she pays a fixed cost of $c \in [0, 1]$. She also receives a benefit of 1 whenever the applicant has *High* quality. With her approval, the game ends; the applicant stops and the remaining evaluators walk away with a payoff of 0. On the other hand, if she rejects the applicant, she receives a payoff of 0. The applicant then proceeds to visit the remaining evaluators, unless none remain.

At the outset of the game, the applicant and all n evaluators share the *prior belief* that the applicant has *High* quality with probability $\rho \in (0, 1)$. Moreover, the evaluators commonly believe that they are equally likely to be anywhere in the applicant's visit order; i.e. that $\mathbb{P}(\tau(k) = i) = \frac{1}{n}$ for all $k, i \in \{1, 2, ..., n\}$. No further information about the applicant's visits is disclosed to evaluators; neither the order $\tau(.)$ he follows, nor the number of evaluators who rejected him already.

Crucially however, an evaluator receiving a visit understands that the applicant was rejected in all his previous visits, however many there might have been. With this information the applicant's visit conveys, she updates her prior belief ρ about the applicant's quality to an *interim belief* ψ. She uses Bayes' Rule to do so.

Dan: update rule clarified (but not in eqm desc.).

Subsequently, she costlessly and privately observes the outcome of a Blackwell experiment $\mathcal{E} = (\mathbf{S}, p_L, p_H)$. The outcome s of this experiment, the evaluator's *signal*, is an element of the finite set $\mathbf{S}$, and has a distribution p_θ over this set given the applicant's quality. Conditional on the applicant's quality, different evaluators' signals are IID. After she observes the signal $s \in \mathbf{S}$, the evaluator updates her beleif about the applicant's quality again; now from the interim belief ψ to a *posterior belief* $\mathbb{P}_\psi(\theta = H \mid S = s)$. For this update too, she uses Bayes' Rule.

Dan: changed. also no longer "signal realisations", just "signal". thoughts?

An evaluator's strategy $\sigma : \mathbf{S} \to [0, 1]$ prescribes a probability of approval $\sigma(s)$ to every possible signal $s \in \mathbf{S}$ she might observe. I call a strategy σ *optimal against the evaluator's interim belief* ψ if, given this interim belief ψ, it maximises her expected payoff. Under such a strategy σ, the evaluator approves after any signal $s \in \mathbf{S}$ which raises her posterior belief that the applicant has *High* quality above c. Likewise, she rejects whenever this posterior belief sinks below c:

$$\sigma(s) = \begin{cases} 0 & \mathbb{P}_\psi(\theta = H \mid S = s) < c \\ 1 & \mathbb{P}_\psi(\theta = H \mid S = s) > c \end{cases}$$

A signal that sets her posterior belief exactly equal to c leaves her indifferent between approving and rejecting the applicant. Any approval probability her strategy dictates after such a signal realisation is consistent with its optimality.

I focus on the *symmetric Bayesian Nash Equilibria* of this game. Hereafter, I reserve the word "equilibrium" for such equilibria unless I state otherwise. A strategy and belief pair (σ^*, ψ^*) is an *equilibrium* of this game if and only if it satisfies the two conditions below:

1. The interim belief ψ^* is *consistent* with the strategy σ^*. That is, an evaluator receiving a visit forms the interim belief ψ^* given others' strategies are σ^*.
2. The strategy σ^* is optimal given the interim belief ψ^*.

I call any strategy σ^* which constitutes part of an equilibrium an *equilibrium strategy*.

2 Belief Formation and Equilibria

Before her verdict, the evaluator who receives a visit must assess the probability that she faces a *High* quality applicant. Her privately observed signal about this applicant's quality plays a crucial part in this assessment. But she obtains her first piece of information even earlier, through *her mere receipt of the applicant's visit*.

The applicant visits this evaluator only if he was rejected by every evaluator he visited earlier. Any such rejections are themselves bad news about the applicant's quality, as they reveal his past evaluators' negative assessments. No information about the number of these past rejections is disclosed to our evaluator. Nonetheless, she is aware of the adverse selection problem she faces: the likelier her peers are to reject the applicant, the likelier she is to be visited by him. Therefore, she interprets the applicant's mere visit as bad news about his quality already.

In particular, when all her peers have the strategies σ, our evaluator understands that an applicant with quality θ faces a probability $r_\theta(\sigma; \mathcal{E})$ of getting rejected from any of his visits, where this probability is given by:

$$r_\theta(\sigma; \mathcal{E}) = 1 - \sum_{j=1}^{m} p_\theta(s_j)\sigma(s_j)$$

She – ex-ante – believes she is equally likely to be anywhere in the applicant's visit order $\tau(.)$. So, she believes that an applicant with quality θ will visit her with probability $\nu_\theta\left(\sigma; \mathcal{E}\right)$ before any of her peers approves him:

$$\nu_\theta\left(\sigma; \mathcal{E}\right) = \frac{1}{n} \times \sum_{k=1}^{n} r_\theta(\sigma; \mathcal{E})^{k-1}$$

Our evaluator's *interim belief* ψ that the applicant who visited her has *High* quality must be *consistent* with these beliefs she holds. Through Bayes Rule, this consistency requirement pins her interim belief down uniquely:

$$\begin{aligned}\psi := \mathbb{P}\left(\theta = H \mid \text{visit received}\right) &= \frac{\mathbb{P}\left(\text{visit received} \mid \theta = H\right) \times \mathbb{P}(\theta = H)}{\mathbb{P}\left(\text{visit received}\right)} \\ &= \frac{\rho \times \nu_H\left(\sigma; \mathcal{E}\right)}{\rho \times \nu_H\left(\sigma; \mathcal{E}\right) + (1-\rho) \times \nu_L\left(\sigma; \mathcal{E}\right)}\end{aligned}$$

After the evaluator updates her prior belief to this interim belief, she observes her private signal $s \in \mathbf{S}$. From this signal, she distils further information about the applicant's quality and updates her interim belief ψ to a final *posterior belief* $\mathbb{P}_\psi\left(\theta = H \mid s\right)$:

$$\mathbb{P}_\psi\left(\theta = H \mid S = s\right) = \frac{\psi \times p_H(s)}{\psi \times p_H(s) + (1-\psi) \times p_L(s)}$$

The information packed in the signal $s \in \mathbf{S}$ is determined exclusively by the conditional probabilities of this outcome, $p_H(s)$ and $p_L(s)$. So for notational convenience, I label every signal $s \in \mathbf{S}$ after a ratio of these conditional probabilities:

$$s = \frac{p_H(s)}{p_H(s) + p_L(s)}$$

I call this ratio the *normalised belief* signal s induces. I merge signals with equal normalised beliefs as there is no reason to distinguish between them. Likewise without loss, I enumerate every signal the evaluator might observe, $\mathbf{S} = \{s_1, s_2, ..., s_m\}$, with indices strictly increasing in the normalised beliefs they induce; $s_1 < s_2 < ... < s_m$.

Using this notation, we can re-express the evaluator's posterior belief upon observing the signal $s_j \in \mathbf{S}$ as simply:

$$\mathbb{P}_\psi\left(\theta = H \mid s_j\right) = \frac{\psi \times s_j}{\psi \times s_j + (1-\psi) \times (1-s_j)}$$

As this expression clarifies, an evaluator's posterior belief is also increasing in the index of the signal $s_i \in \mathbf{S}$ she observes: it is a simple rescaling of the *normalised belief* this signal induces with the evaluator's interim belief. Note that the normalised belief equals the posterior belief when the evaluator's interim belief assigns equal probability to either quality; $\psi = 0.5$.

Whenever a strategy σ^* is optimal against the unique interim belief ψ^* consistent with it, the pair (σ^*, ψ^*) forms an equilibrium. In principle, there might be many such pairs, or none. I set the ground in Proposition 1 by ruling this last possibility out: an equilibrium is always guaranteed to exist. Also in Proposition 1, I describe some properties of these equilibria that are fundamental to the rest of our analysis.

Proposition 1. Where Σ is the set of evaluators' equilibrium strategies:

1. Σ is non-empty and compact.

2. Any equilibrium strategy σ^* is *monotone*: $\sigma^*(s) > 0$ for some $s \in \mathbf{S}$ implies that $\sigma^*(s') = 1$ for every $s' \in \mathbf{S}'$ such that $s' > s$.

3. All equilibria exhibit adverse selection: $\psi^* \leq \rho$ for any interim belief ψ^* consistent with an equilibrium strategy $\sigma^* \in \Sigma$.

I relegate the full proof of Proposition 1 to Section 8. Instead, I discuss its proof in broad strokes here. To establish the existence of an equilibrium, I construct a *best response correspondence* Φ for evaluators. Φ maps any strategy σ to the set of all strategies optimal against the unique interim belief consistent with σ. Put differently, $\Phi(\sigma)$ gives the set of strategies maximising an evaluator's expected payoff when all her peers use the strategy σ. Note that a strategy σ^* is an equilibrium strategy if and only if it is a fixed point of this best response correspondence; $\sigma^* \in \Phi(\sigma^*)$. I show that Φ indeed has a fixed point, by a routine application of Kakutani's Fixed Point Theorem. To this end, I prove that Φ is upper semi-continuous, which establishes the compactness of the set of equilibrium strategies as well.

Monotonicity is a simple consequence of optimality. Higher signals induce higher posterior beliefs; so if an evaluator (weakly) prefers approving her applicant after a signal, she (strictly) prefers it after a higher one too. A crucial consequence of monotonicity is that a *Low* quality applicant is likelier to be rejected after any of his visits, as his evaluators are likelier to observe lower signals. Thus, each evaluator risks getting visited by an applicant who was *adversely selected* through his past rejections. This pushes their interim beliefs ψ below their prior ρ.

Though an equilibrium is guaranteed to exist, it need not be unique. I illustrate this with a simple example which I will revisit on occasion throughout this paper. Consider two evaluators who have the prior belief $\rho = 0.5$ about the applicant's quality, and a cost $c = 0.2$ of approving him. The experiment $\mathcal{E}$ whose outcome they observe is binary; $\mathbf{S} = \{0.2, 0.8\}$. This outcome has the distribution:

$$p_L(s) = \begin{cases} 0.8 & s = 0.2 \\ 0.2 & s = 0.8 \end{cases} \qquad\qquad p_H(s) = \begin{cases} 0.2 & s = 0.2 \\ 0.8 & s = 0.8 \end{cases}$$

One equilibrium strategy for evaluators in this example is to approve every applicant who visits them. Doing so eliminates adverse selection: evaluators never receive an applicant with a past rejection, so their interim belief ψ always equals their prior $\rho = 0.5$. However, at this interim belief, the low signal $s = 0.2$ still implies a 20% probability that the applicant has *High* quality, rendering his approval against the cost 0.2 optimal.

This equilibrium, however, is not unique. There is yet another equilibrium where evaluators approve the applicant only upon the high signal, $s = 0.8$. Their selectivity triggers adverse

Removed totally orderedness to proposition 2 where it's more natural to discuss it.

Also changed eqm. existence proof. Previously was a direct proof with algorithm using IVT. Now Kakutani FPT.

Dan: I changed this paragraph entirely.

selection: each evaluator risks being visited by a past reject, thus revises her interim belief ψ below her prior $\rho = 0.5$:

$$\psi = \frac{1 + 0.2}{(1 + 0.2) + (1 + 0.8)} = 0.4$$

Consequently, she places a 1/7 probability on the applicant having *High* quality upon observing the low signal $s = 0.2$; justifying his rejection. She still finds approving the applicant optimal upon the high signal $s = 0.8$, though. Even at this interim belief, she places a probability greater than 70% on him having *High* quality when she observes this signal.

In this latter equilibrium, an applicant – regardless of his quality – faces higher rejection chances in any of his visits. His evaluators are *more selective*; any signal they might observe leads to a (weakly) higher chance of rejection. It is natural in general to try and compare equilibria in their *selectivity*, whenever we face multiple.

Definition 1. Where σ' and σ are two strategies for evaluators, σ' is *more selective than* σ (or, σ is *less selective than* σ') if $\sigma'(s) \leq \sigma(s)$ for all $s \in \mathbf{S}$.

While natural, the *selectivity* (or *pointwise*) order might initially appear restrictive. This impression is misleading. In fact, the set of equilibrium strategies is *totally ordered* (or, a *chain*) under this order; any two equilibrium strategies can be compared under it. Furthermore, this set has both a *most selective* and *least selective* element, marking its extremes. I refer to them as the *extreme equilibria* in the sequel. The two equilibria we identified in our example earlier were, in fact, its extreme equilibria.

Lemma 1. The set of evaluators' equilibrium strategies Σ is *totally ordered* under the selectivity order. Moreover, Σ contains a *most selective* and *least selective* strategy, $\hat{\sigma} \in \Sigma$ and $\check{\sigma} \in \Sigma$ respectively, such that:

$$\hat{\sigma}(s) \leq \sigma^*(s) \leq \check{\sigma}(s) \qquad \text{for all } s \in \mathbf{S}$$

Proof. By Proposition 1, the set of equilibrium strategies Σ is a subset of the set of monotone strategies. The latter is a chain under the *selectivity* order; for two monotone strategies σ and σ', we have:

$$\sigma'(s_j) > \sigma(s_j) \implies 1 = \sigma'(s_{j'}) \geq \sigma(s_{j'}) \qquad \text{for any } s_{j'} > s_j \in \mathbf{S}$$

Since any subset of a chain is also a chain, Σ is a chain too.

By Proposition 1, Σ is a compact set. Since it is also a chain, by a suitably general Extreme Value Theorem (see, for instance, Theorem 27.4 in Munkres, 2000) it has a *minimum* and *maximum* element, $\hat{\sigma}$ and $\check{\sigma} \in \Sigma$ respectively, with respect to this order:

$$\hat{\sigma}(s) \leq \sigma^*(s) \leq \check{\sigma}(s) \qquad \text{for all } s \in \mathbf{S}$$

$\square$

The applicant – regardless of his quality – is worse off with more selective evaluators. His evaluators grow more reluctant to approve him after any signal they might observe, so he faces a higher rejection risk in any of his visits. How moving to more selective equilibria affects evaluators' payoffs is less clear. Their payoffs are determined by how they balance their two key objectives: identifying and approving a *High* quality applicant, and rejecting a *Low* quality one. The expression $\Pi(\sigma;\mathcal{E})$, the *sum* of evaluators' payoffs when each use the strategy σ, highlights this:

$$\begin{aligned}\Pi(\sigma;\mathcal{E}) := \rho \times (1-c) \times \mathbb{P}\,(\text{some ev. approves when all use strategies } \sigma \mid \theta = H) \\ -(1-\rho) \times c \times [1 - \mathbb{P}\,(\text{all ev.s reject when all use strategies } \sigma \mid \theta = L)]\end{aligned} \tag{2.1}$$

Each evaluator expects simply $\left(\frac{1}{n}\right)^{\text{th}}$ of this sum of course, as the equilibrium is symmetric.

Increased selectivity has counteracting effects on these two objectives. It mitigates their risk of approving a *Low* quality applicant when they face one, sparing them a cost of c. However, this comes at the expense of curbing the approval chances of a *High* quality applicant too, which means forsaking a payoff of $1-c$. In principle, increased selectivity can therefore both be a vice and a virtue.

Our previous example, where we had identified two equilibria, illustrates these competing effects of increased selectivity. In the least selective equilibrium, evaluators approve all applicants; either *High* quality, or *Low*. Their payoffs, therefore, sum to:

$$\Pi(\check{\sigma};\mathcal{E}) = 0.5 \times [(1-c) - c] = 0.3$$

In the most selective equilibrium on the other hand, an evaluator rejects an applicant for whom she observes the low signal $s = 0.2$. This depresses the approval chances of any applicant. A *High* quality applicant faces a probability $p_H^2(0.2) = 0.04$ of getting rejected by both his evaluators. This probability is higher for a *Low* quality applicant, $p_L^2(c) = 0.64$. Under these more selective strategies, evaluators' payoffs sum to:

$$\Pi(\hat{\sigma};\mathcal{E}) = 0.5 \times \left[(1-c) \times (1 - p_H(0.2)^2) - c \times \left(1 - p_L(0.2)^2\right)\right] = 0.348$$

Despite reducing the approval chances of both *Low* and *High* quality applicants, selectivity pays off for our evaluators in this example.

Why increased selectivity ends up helping evaluators is clear in this example. When they switch to the more selective equilibrium, evaluators push out only applicants for whom *both* of them saw low signals. The probability that such an applicant has *High* quality is less than 6%. By rejecting him, they save a cost of 0.2 against an expected benefit of 0.06, which raises their payoffs.

In more intricate examples, it becomes less clear whether evaluators will benefit from increased selectivity; which can lead them to reject desirable applicants. Consider, for instance, a richer experiment $\mathcal{E}'$ with three possible outcomes $\mathbf{S}' = \{0.2, 0.4, 0.8\}$ for the evaluators in our running example. The outcome of $\mathcal{E}'$ has the distribution:

$$p'_L(s) = \begin{cases} 0.48 & s = 0.2 \\ 0.36 & s = 0.4 \\ 0.16 & s = 0.8 \end{cases} \qquad p'_H(s) = \begin{cases} 0.12 & s = 0.2 \\ 0.24 & s = 0.4 \\ 0.64 & s = 0.8 \end{cases}$$

Again, evaluators treat the signal $s = 0.4$ differently in different equilibria. Both σ and σ' defined below are equilibrium strategies. In the former, more selective equilibrium, evaluators reject upon the signal $s = 0.4$. In the latter less selective one, they approve[1] .

$$\sigma(s) := \begin{cases} 0 & s \in \{0.2, 0.4\} \\ 1 & s = 0.8 \end{cases} \qquad \sigma'(s) := \begin{cases} 0 & s = 0.2 \\ 1 & s \in \{0.4, 0.8\} \end{cases}$$

It is no longer as clear whether switching from the less selective strategies σ' to the more selective σ will benefit evaluators. As an unpleasant consequence of this switch, an applicant for whom both evaluators observe the signal $s = 0.4$ gets rejected. The probability that he has *High* quality exceeds 0.3; so his approval against the cost $c = 0.2$ would have benefited evaluators. On the other hand, consider another applicant approved unknowingly by his second evaluator upon the signal $s = 0.4$, despite his first rejection upon $s = 0.2$. He has a probability less than 0.15 of having *High* quality. Upon this switch, evaluators successfully push this applicant out too, to their benefit.

should I detail what interim belief they induce, and verify that this is an eqm?

Nonetheless, if we calculated evaluators' payoffs under both equilibria, we would find that the overall effect is driven by this latter applicant; their increased selectivity – yet again – leaves evaluators better off. The conclusion these examples share in fact demonstrates a general phenomenon I establish in Proposition 2. Evaluators' trade-off between more and less selective equilibria is *always* resolved in favour of the former. Notably, this brings the welfare of the applicant, unambiguously harmed by selectivity, into conflict with the evaluators'.

Proposition 2. Where σ^* and σ^{**} are two equilibrium strategies such that σ^{**} is more selective than σ^*, evaluators' expected payoffs under σ^{**} exceed those under σ^*; $\Pi(\sigma^{**}; \mathcal{E}) \geq \Pi(\sigma^*; \mathcal{E})$.

Selectivity thus offers a very powerful comparison between equilibria, besides a very natural

[1]It is easy to verify that these are, indeed, equilibrium strategies. They give rise to the respective interim beliefs $\psi = \frac{1+0.24+0.12}{(1+0.24+0.12)+(1+0.36+0.48)} = 0.425$ and $\psi' = \frac{1+0.12}{(1+0.12)+(1+0.48)} =$

one. We can use it to compare any two equilibria, and to determine both the applicant's and evaluators' relative welfare. *Extreme* equilibria deserve particular focus. The *most selective* equilibrium maximises evaluators payoffs across all equilibria while minimising the applicant's approval chances. The *least selective* equilibrium, vice versa. They remain under my spotlight in the remainder of this paper.

Proposition 2 follows as a corollary to Lemma 2, which establishes that in fact deviating to *any* less selective strategy hurts evaluators' payoffs when they start from an equilibrium strategy.

Lemma 2. Let σ^* and σ be two monotone strategies, where σ^* is more selective than σ. If σ^* is an equilibrium strategy, then $\Pi(\sigma^*; \mathcal{E}) \geq \Pi(\sigma; \mathcal{E})$.

Proof. See Section 8. □

Besides birthing Proposition 2, Lemma 2 highlights an important contrast between the problem a *single* evaluator with no peers faces, and the one *multiple* evaluators do. A single evaluator with no peers faces no adverse selection; no evaluator could have received her applicant earlier. So her interim belief *is* her prior belief, regardless of the strategy she chooses. In equilibrium, her strategy must be optimal against this belief. This equilibrium strategy is unique – up to how she breaks ties when indifferent. Any equilibrium strategy gives her the same expected payoff, and deviation can leave her only worse off.

With multiple evaluators, a non-trivial multiplicity of equilibria becomes possible. There might be multiple strategies which motivate an interim belief they are optimal against. Crucially, evaluators' payoffs vary between these different equilibria; joint deviations to *more selective* equilibria benefit them. Lemma 2 highlights that only increased selectivity might pay off, though. Deviating – individually or jointly – from any equilibrium to a *less selective strategy* still hurts an evaluator's payoffs, just as it would had she no peers.

As the earlier examples illustrate, two kinds of applicants are pushed out when evaluators move to more selective equilibria. Some fall through the cracks: all evaluators reject them, although none have convincing evidence that they have *Low* quality. Other applicants, on the other hand, face decisive rejections from most evaluators, and – once they become more selective – lose the benefit of doubt a few unsuspecting evaluators would grant them. I prove Lemma 2 by showing that the overall effect is always driven by the latter group of applicants, whose rejection benefits evaluators.

3 More Informative Experiments and Equilibrium Payoffs

As we discussed in the previous section, evaluators' payoffs are determined fundamentally by how well they can distinguish and approve a *High* quality applicant while rejecting a *Low* quality one. The information they obtain about the applicant's quality from their experiment $\mathcal{E}$ lies at the heart of this exercise. One might, perhaps naturally, hypothesise that a *more informative* experiment is the key to improving evaluators' welfare. Eventually, more information helps evaluators identify their applicant's quality better. This ought to ease the tension between their two fundamental objectives, allowing them to boost their payoffs.

This hypothesis would indeed be correct if we only had *one* evaluator. With no peers, she faces a simple decision problem: given her fixed prior belief about her applicant's quality, she must choose whether to approve him given the outcome of her experiment $\mathcal{E}$. As Blackwell's seminal result (1953) establishes, observing instead the outcome of an experiment $\mathcal{E}'$ that is *(Blackwell) more informative* than $\mathcal{E}$ would indeed leave her better off. In fact – absent further particulars about her decision problem (namely, her approval cost c and prior belief ρ) – only a Blackwell more informative experiment can guarantee her a higher expected payoff[2]. This is precisely because more information relaxes our evaluator's key trade-off. She can reject *Low* quality applicants more frequently without compromising *High* quality applicants more often, and (or) vice versa [3]. I illustrate this in Figure 1 in the next subsection, in the context of *binary* experiments (where $\mathbf{S} = \{s_1, s_2\}$).

Nevertheless, this naïve hypothesis fails in our current setting. Re-expressing an individual evaluator i's payoff $\pi_i(\sigma; \mathcal{E})$ showcases what goes wrong:

$$\begin{aligned}\pi_i(\sigma;\mathcal{E}) = {} & \mathbb{P}\left(\text{applicant visits } i\right) \\ & \times \Big[\psi \times (1-c) \times \mathbb{P}\left(i \text{ approves} \mid \theta = H\right) + (1-\psi) \times (-c) \times \mathbb{P}\left(i \text{ approves} \mid \theta = L\right)\Big]\end{aligned}$$

Evaluator i's payoff, like others', is determined by how well she can tailor her decisions to the applicant's quality. However, she is also affected by the extent of *adverse selection* she faces. The applicant might not visit her at all, and if he does, he might be very unlikely to have *High* quality given no other evaluator approved him so far.

The extent of adverse selection evaluator i faces is shaped by *other* evaluators' strategies. These evaluators do not account for the adverse selection their decisions' impose on her; just as she disregards the adverse selection she imposes on them. This adverse selection externality can transform more information into a threat. By Pushing evaluators to improve their selection quality, more information might accentuate the adverse selection they impose on each other. This might eclipse their improved ability to evaluate the applicant, and leave all evaluators

[2]Blackwell's Theorem (1953) is in fact weaker. It states that an experiment $\mathcal{E}'$ offers the decision maker a higher expected payoff than $\mathcal{E}$ *regardless of the decision problem she faces* if and only if $\mathcal{E}'$ is (Blackwell) more informative than $\mathcal{E}$. This establishes that (Blackwell) more informativeness is *sufficient* to secure our evaluator a higher expected payoff. However, it does not establish its *necessity*; our evaluator does not just face *any* decision problem, but a two state - two action one. Nonetheless, Blackwell's Theorem retains its sufficiency for this class of decision problems too. I present a self contained proof for this in Section 8.2, Lemma 4 for completeness.

[3]See Blackwell and Girshick, 1954's Theorems 12.2.2 and 12.4.2 for a textbook exposition of these classic results.

worse off.

thoughts on this par?

Our simple example illustrates this possibility. To simplify this illustration, I take evaluators' approval cost to be $c = 0.5$ instead. The least selective equilibrium earlier does not survive this increase in the approval cost; even without adverse selection, evaluators find it optimal to reject their applicant upon the low signal $s = 0.2$. In fact, the most selective equilibrium we identified earlier now becomes the unique one: evaluators approve the applicant upon the high signal $s = 0.8$, but reject him upon the low signal, $s = 0.2$. With this modified cost, evaluators' payoff in this equilibrium becomes:

$$0.5 \times \left[(1-c) \times (1 - p_H(0.2)^2) - c \times \left(1 - p_L(0.2)^2\right)\right] \Big|_{c=0.5} = 0.15$$

Now, consider swapping our evaluators' experiment $\mathcal{E}$ with a more informative binary experiment $\mathcal{E}^g$ whose possible outcomes are in the set $\mathbf{S}^g = \{c, 1\}$. The high signal in $\mathcal{E}^g$ carries *conclusive good news*; since no evaluator observes it otherwise, any evaluator to observe it definitively concludes that the applicant has *High* quality. The evidence the low signal carries for *Low* quality, however, is no stronger than before. For the *same* interim belief, observing the low signal from $\mathcal{E}^g$ leads to the same posterior belief as observing the low signal from $\mathcal{E}$ does. The outcomes of experiment $\mathcal{E}^g$ have the distribution:

$$p_L^g(s) = \begin{cases} 1 & s = 0.2 \\ 0 & s = 1 \end{cases} \qquad p_H^g(s) = \begin{cases} 0.25 & s = 0.2 \\ 0.75 & s = 1 \end{cases}$$

Under this more informative experiment, the unique equilibrium remains the one where evaluators approve upon the high signal $s = 1$, but reject upon the low signal $s = 0.2$. Thus, evaluators manage to avoid approving *any Low* quality applicant, albeit forsaking *High* quality applicants more often than earlier. This improvement in their information pays off; evaluators' payoffs now sum up to:

$$0.5 \times \left[(1-c) \times \left(1 - p_H^g(0.2)^2\right)\right] \Big|_{c=0.5} \approx 0.23$$

surpassing their payoffs of 0.15 under their original experiment $\mathcal{E}$.

Now consider another binary experiment, $\mathcal{E}^b$, again more informative than evaluators' original experiment $\mathcal{E}$. The possible outcomes of $\mathcal{E}^b$ lie in the set $\mathbf{S}^b = \{0, 0.8\}$ this time, with distribution:

$$p_L^b(s) = \begin{cases} 0.75 & s = 0 \\ 0.25 & s = 0.8 \end{cases} \qquad p_H^b(s) = \begin{cases} 0 & s = 0 \\ 1 & s = 0.8 \end{cases}$$

It is the low signal $s = 0$ which carries *conclusive bad news* this time; any evaluator concludes that the applicant has *Low* quality upon observing it. Would our evaluators benefit if we swapped their experiment not with $\mathcal{E}^g$ but with $\mathcal{E}^b$ instead?

The answer, it turns out, is no this time. Still, evaluators approve upon the high signal $s = 0.8$ and reject upon the low, $s = 0$, in the unique equilibrium. This now guarantees that they *always* approve a *High* quality applicant, though at the expense of approving *Low* quality applicants more often than before. Their payoffs now sum up to:

$$0.5 \times \left[(1-c) - c \times \left(1 - p_L(0)^2\right)\right] \Bigg|_{c=0.5} \approx 0.14$$

falling behind their payoffs under the original experiment $\mathcal{E}$.

The pattern in this example is a general one: while *stronger good news*, in the appropriate sense, always benefit evaluators, *stronger bad news* eventually hurts them. In the remainder of this section, I uncover and explore this general pattern. I start by restricting evaluators to binary experiments in Section 3.1. Theorem 1 there characterises how giving evaluators a Blackwell more informative binary experiment affects their payoffs, and discusses the intuition behind this. The general result I present in Theorem 2 in Section 3.2 builds on this groundwork. It characterises how an *arbitrary* Blackwell improvement of evaluators' experiment affects their payoffs, and shows the intuition laid out by Theorem 2 fully generalises. Throughout, I focus on evaluators' payoffs across *extreme equilibria*. They delineate the boundaries of both the evaluators' and the applicant's welfare across equilibria, hence command the highest importance.

is this the right language?

3.1 Evaluators with Binary Experiments

In this section, I consider evaluators who observe the outcome of a *binary* experiment. Such an experiment has two possible outcomes, $\mathbf{S} = \{s_1, s_2\}$, which I respectively rename $\mathbf{S} = \{s_L, s_H\}$ for notational convenience. The low outcome induces the normalised belief $s_L \in [0, 0.5]$, carrying *bad news* about the applicant's quality. In contrast, the high outcome induces the normalised belief $s_H \in [0.5, 1]$, carrying *good news* about the applicant's quality.

How do evaluators' equilibrium payoffs change when we swap this experiment with a more informative binary experiment $\mathcal{E}'$? Here, I answer this question. This answer also lays the building block and key intuition for the next section, where I ask how an *arbitrary* Blackwell improvement affects evaluators' payoffs.

Where $\mathcal{E}$ and $\mathcal{E}'$ are two experiments, with possible outcomes in $\mathbf{S} = \{s_L, s_H\}$ and $\mathbf{S}' = \{s'_L, s'_H\}$ respectively, $\mathcal{E}'$ is *(Blackwell) more informative than* (or *(Blackwell) improves on*) $\mathcal{E}$ if and only if it carries both *stronger good news* and *stronger bad news* than $\mathcal{E}$[4]; i.e.:

$$s'_L \leq s_L \qquad\qquad s'_H \geq s_H$$

[4]See Section 12.5 in Blackwell and Girshick, 1954 for a textbook exposition of this classic result.

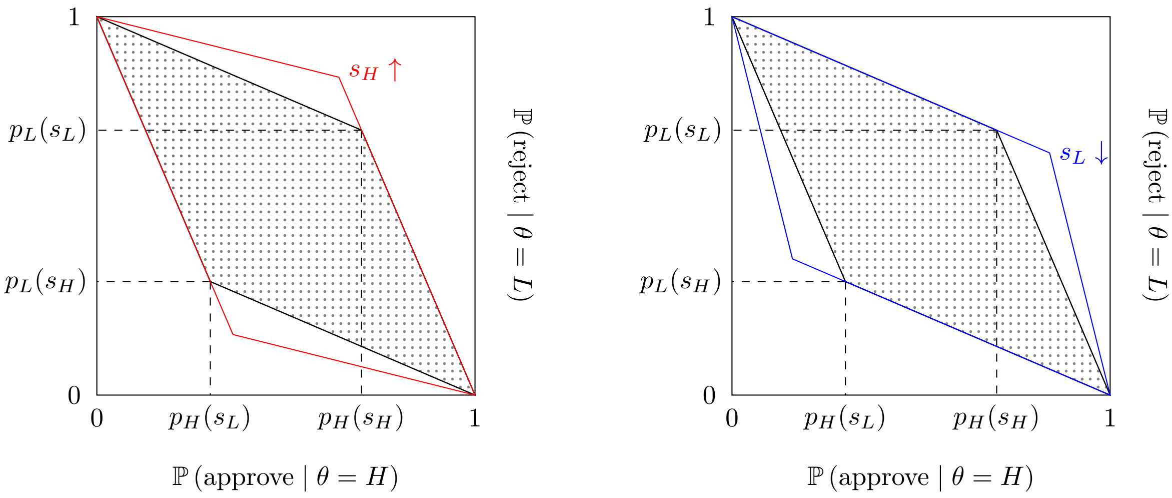
For any fixed interim belief $\psi \in [0,1]$, the experiment $\mathcal{E}'$ helps an evaluator form more confident assessments of the applicant's quality. Its low outcome leaves her more confident that her applicant has *Low* quality, and its high outcome that he has *High*:

$$\mathbb{P}_\psi\left(\theta = H \mid s'_L\right) \leq \mathbb{P}_\psi\left(\theta = H \mid s_L\right) \qquad \mathbb{P}_\psi\left(\theta = H \mid s'_H\right) \geq \mathbb{P}_\psi\left(\theta = H \mid s_H\right)$$

idk how much i like "its low outcome" blah rather than "signal".

also using "outcome" and "signal" to interchangeably.

Dan: ensuing discussion expanded.



(a) Stronger Good News: $s_H \uparrow$ (b) Stronger Bad News: $s_L \downarrow$

Figure 1: Evaluators' Trade-Offs with More Informative Experiments

Figure 1 illustrates how switching to the more informative experiment $\mathcal{E}'$ transforms evaluators key trade-off. In the unit squares it depicts, each point represents approval probabilities an evaluator *might* be able to achieve with some experiment and strategy. The horizontal axis marks the probability that she approves a *High* quality applicant, and the vertical axis that she rejects a *Low* quality one.

Of course, her experiment $\mathcal{E}$ limits her from implementing some of these approval probabilities. The top right corner of the square, for instance, where an evaluator certainly approves *High* quality applicants and rejects *Low* quality ones, requires $\mathcal{E}$ to *fully reveal* the applicant's quality. Instead, the dotted region depicts the subset of approval probabilities she can implement with the experiment $\mathcal{E}$ by freely varying her strategy. For example, the top corner of that dotted parallelogram corresponds to what she can implement by "approving upon the high signal, and rejecting upon the low".

As the Figure illustrates, this region of implementable approval probabilities expands with an experiment $\mathcal{E}'$ carrying *stronger good news* (left panel) or *stronger bad news* (right panel). With more information, the evaluator can choose from a strictly larger set of approval probabilities. A *single* evaluator can never suffer from this expansion. However, she might if she is surrounded by other evaluators who enjoy the same expansion. Their choices determine the extent of adverse selection she is exposed to, and the new choices they make might aggravate it.

How does, then, more information affect evaluators' *equilibrium* payoffs? As the preceding discussion foreshadowed, Theorem 1 reveals that this hinges precisely on whether evaluators receive *stronger good news* or *stronger bad news*. Stronger good news – higher s_H – always increase evaluators' payoffs, both in the *most selective* equilibria and the *least*. The effect of stronger bad news – lower s_L – in contrast, is more delicate. Initially, evaluators benefit from stronger bad news. However, once s_L falls below a cutoff, their payoffs fall as bad news get even stronger. This effect occurs across both the most and the least extreme equilibria; though the cutoff for these two equilibria may differ.

Theorem 1. Let the experiment $\mathcal{E}$ be binary with outcomes inducing the normalised beliefs $s_L \in [0, 0.5]$ and $s_H \in [0.5, 1]$. Increasing s_H weakly increases evaluators' payoffs across the extreme equilibria. In contrast, as s_L decreases, evaluators' payoffs in the most (least) selective equilibrium:

1. weakly improve as long as s_L remains below a cutoff $\hat{s_L}$ ($\check{s_L}$),

2. weakly decrease once s_L falls below this cutoff.

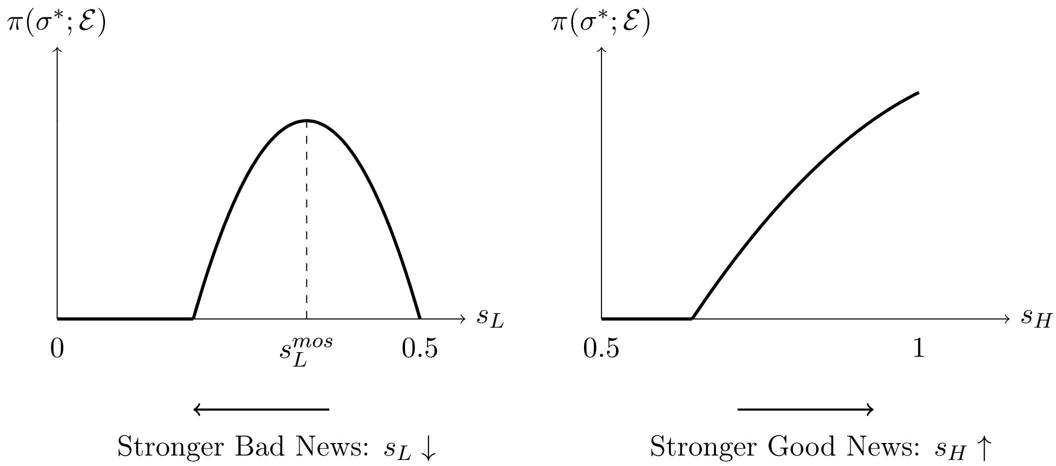


Figure 2: Theorem 1 illustrated

Figure 2 illustrates Theorem 1 for an example where evaluators' approval cost c weakly exceeds their prior belief ρ. This conveniently forces a unique equilibrium across experiments, whose strategies I indiscriminately label σ^*. Examples with multiple equilibria would produce qualitatively the same plots; both for the most selective equilibria different experiments induce, and the least.

In the ensuing discussion, I explain the discrepancy between the effect of strengthening good news and strengthening bad news on evaluators' payoffs. I relegate the full proof of Theorem 1 to Section 8, which builds on the forces I lay out here. The eager reader will notice that Theorem 1 does not elaborate on the cutoff below which stronger bad news harm evaluators' payoffs. As I clarify the forces driving Theorem 1, I cast light on this cutoff as well. Building on that discussion, I characterise this cutoff in Proposition 3.

Carlos: I actually have a nice result here. I can prove it using another one in the appendix.
"Equilibrium adverse selection always increases with the informativeness of the signal".
This is why the reader should care about what's to come next anyway: because one's first instinct towards a proof will be fruitless. I don't know how to insert this yet. It yields another nice result:
"The expected quality of an approved applicant falls when s_L falls, in equilibrium."

Any change in the evaluators' experiment ultimately affects who they approve and who they reject. A *High* quality applicant they would otherwise reject might this time receive an evaluator's approval, or a *Low* quality applicant who would otherwise slip through their net might instead face a wall of rejections. Other changes might be less welcome: evaluators might inadvertently dismiss more *High* quality applicants, or fail to do so with *Low* quality ones. It is the applicants they affect who determine how stronger good news or stronger bad news influence evaluators' payoffs.

Let us start with the case of *stronger bad news.* Consider switching evaluators' experiment from $\mathcal{E}$ to $\mathcal{E}'$, whose possible outcomes $\mathbf{S}' = \{s'_L, s'_H\}$ induce the normalised beliefs:

$$s'_L = \frac{p'_H(s'_L)}{p'_H(s'_L) + p'_L(s'_L)} = s_L - \delta \qquad\qquad s'_H = \frac{p'_H(s'_H)}{p'_H(s'_H) + p'_L(s'_H)} = s_H$$

for some small $\delta > 0$. Experiment $\mathcal{E}'$ thus offers *marginally stronger bad news* than $\mathcal{E}$, but the same the strength of good news as the latter. Evaluators' equilibrium strategies will of course react to this switch. Nevertheless, ignoring this strategic response allows for a clearer intuition. So instead, let us simply assume that under both experiments, evaluators approve whenever they observe the "high" outcome, and reject whenever they observe the "low" outcome. Which applicants' outcomes does the switch from $\mathcal{E}$ to $\mathcal{E}'$ affect?

The clearest way to answer this question is by reinterpreting this improvement in evaluators' information. Imagine, instead of replacing their original experiment $\mathcal{E}$ wholesale, that evaluators observe an *auxiliary signal* $\hat{S}$ in addition to their original S. We will construct this auxiliary signal $\hat{S}$ carefully so that it completes the information evaluators garner from their original experiment $\mathcal{E}$ to the one they could from $\mathcal{E}'$. I illustrate this construction in Figure 3. The reader might benefit from referring to it throughout the ensuing discussion.

This auxiliary signal $\hat{S}$ evaluators observe is also binary, with possible realisations $\hat{s} \in \{\hat{s_L}, \hat{s_H}\}$. Conditional on the applicant's quality θ, its outcome is independent both from the evaluator's original signal S and anything other evaluators observe. It has a distribution:

$$\hat{p_H}(\hat{s_H}) = \varepsilon \times \frac{s_H}{1 - s_H} \qquad\qquad \hat{p_L}(\hat{s_H}) = \varepsilon \times \frac{s_L}{1 - s_L}$$

ε, like δ, is a small positive number. It is related intimately to δ, as I explain shortly.

An evaluator observes the realisation of this auxiliary signal $\hat{s}$ *only* if the initial signal she observes is low, $s = s_L$. If she observes $\hat{s} = \hat{s_H}$ following this initial low signal, her belief that the applicant has *High* quality jumps to what it would be had she observed $s = s_H$ straightaway. This is most visible from the likelihood ratio for this signal pair:

$$\frac{\mathbb{P}\left(s = s_L, \hat{s} = \hat{s_H} \mid \theta = H\right)}{\mathbb{P}\left(s = s_L, \hat{s} = \hat{s_H} \mid \theta = L\right)} \quad = \quad \frac{p_H(s_L)}{p_L(s_L)} \times \frac{\hat{p_H}(\hat{s_H})}{\hat{p_L}(\hat{s_H})} \quad = \quad \frac{s_L}{1 - s_L} \times \frac{\frac{s_H}{1-s_H}}{\frac{s_L}{1-s_L}} \quad = \quad \frac{s_H}{1 - s_H}$$

If she instead observes $\hat{s} = \hat{s_L}$ though, she grows yet more confident that the applicant has *Low* quality. Again note this from the likelihood ratio for this signal pair, labelled $(L, \hat{L})$:

$$\frac{\mathbb{P}\left(s = s_L, \hat{s} = \hat{s_L} \mid \theta = H\right)}{\mathbb{P}\left(s = s_L, \hat{s} = \hat{s_L} \mid \theta = L\right)} \quad = \quad \frac{p_H(s_L)}{p_L(s_L)} \times \frac{\hat{p_H}(\hat{s_L})}{\hat{p_L}(\hat{s_L})} \quad = \quad \underbrace{\frac{s_L}{1 - s_L} \times \frac{1 - \frac{s_H}{1-s_H} \times \varepsilon}{1 - \frac{s_L}{1-s_L} \times \varepsilon}}_{(L,\hat{L})} \quad < \quad \frac{s_L}{1 - s_L}$$

The likelihood ratio $(L, \hat{L})$ decreases continuously and monotonically as ε rises from 0 to $\frac{1-s_H}{s_H}$. We can thus choose ε so that this likelihood ratio equals the one for the low outcome of experiment $\mathcal{E}'$, $s' = s'_L$. This latter likelihood ratio is labelled (L') below:

it is a confusing jump from likelihood ratios to "normalised beliefs"?

$$\frac{\mathbb{P}\left(s' = s'_L \mid \theta = H\right)}{\mathbb{P}\left(s' = s'_L \mid \theta = L\right)} \quad = \quad \frac{p'_H(s'_L)}{p'_L(s_L)} \quad = \quad \underbrace{\frac{s_L - \delta}{1 - (s_L - \delta)}}_{(L')}$$

Note that the value of ε equating these likelihood ratios is a continuous and strictly increasing function of δ.

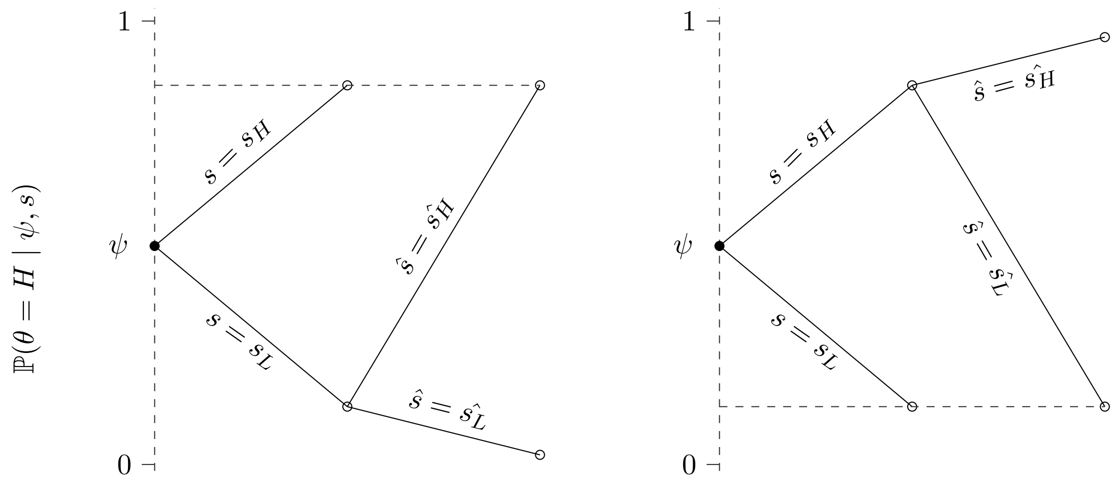
When the likelihood ratios $(L, \hat{L})$ and (L') are equal, the information an evaluator obtains from observing the outcome of $\mathcal{E}'$ is equivalent to the one she does by observing the signal pair $(S, \hat{S})$. Receiving a high signal, either $s = s_H$ or $\hat{s} = \hat{s_H}$, carries the same information as the high outcome $s' = s'_H$ from experiment $\mathcal{E}'$. Receiving only the low signals $s = s_L$ and $\hat{s} = \hat{s_L}$ on the other hand, carries the same information as observing $s' = s'_L$ from $\mathcal{E}'$. Our evaluator can easily replicate the outcome of her original strategy with this signal pair; she approves the applicant if she receives a high signal, but rejects him otherwise.

We can interpret this auxiliary signal as an evaluator's "re-evaluation" of her initial rejection decisions. Her initial approvals remain final; so this re-evaluation does not affect applicants who received some evaluator's approval anyway. Instead, it affects the applicants who were initially rejected by *all* n evaluators. Each of these n re-evaluations might overturn an evaluator's negative verdict, and grant this applicant the approval he seeks. It is this applicant, who some evaluator approves upon her re-evaluation, who drives the change in evaluators' payoffs.

Dan: added last sentence here.

How likely is this applicant to have *High* quality? All information we can harvest about

Dan: changed paragraph



(a) Stronger Bad News

(b) Stronger Good News

Figure 3: Improving Binary Experiments with Auxiliary Signals

this is contained in the signals evaluators would observe, if they *all* were to (re-)evaluate him. What is certain is that initially, they all observed low signals; this led to the applicant's initial rejections. Upon their re-evaluations, how many would see a high signal?

The applicant we are inspecting is approved upon these re-evaluations, so we know *at least one* evaluator must have seen a high signal. Discouragingly, however, she would *almost surely be the only one* to see it, for small δ and therefore ε. Recall from our construction that the probability any evaluator observes the signal $\hat{s} = \hat{s_H}$ for an applicant is proportional to ε; and so the probability that any k evaluators observe it, to ε^k. As ε shrinks to 0, the probability that multiple evaluators observed this signal vanishes in favour of the probability that just one of them did. Hence, that one evaluator who observes the signal $\hat{s} = \hat{s_H}$ approves the applicant against the backdrop of $n-1$ $\hat{s} = \hat{s_L}$ signals her peers observed. The stronger the bad news those signals carry, the less likely he is to have *High* quality. Inspecting his signals' likelihood ratios reveals precisely when bad news are *too strong* for evaluators to benefit from this applicant:

Dan: changes end.

$$
\begin{aligned}
&\lim_{\delta \to 0} \mathbb{P}(\theta = H \mid \ n-1 \ \hat{s} = \hat{s_L} \text{ signals and one } \hat{s} = \hat{s_H}) \geq c \\
\iff &\lim_{\delta \to 0} \frac{\mathbb{P}\left(\theta = H \mid \ n-1 \ \hat{s} = \hat{s_L} \text{ signals and one } \hat{s} = \hat{s_H}\right)}{\mathbb{P}\left(\theta = L \mid \ n-1 \ \hat{s} = \hat{s_L} \text{ signals and one } \hat{s} = \hat{s_H}\right)} \geq \frac{c}{1-c} \\
\iff &\lim_{\delta \to 0} \frac{\rho}{1-\rho} \times \left(\frac{\hat{s_L}}{1-\hat{s_L}}\right)^{n-1} \times \frac{s_H}{1-s_H} \geq \frac{c}{1-c} \\
\iff &\frac{\rho}{1-\rho} \times \underbrace{\left(\frac{s_L}{1-s_L}\right)^{n-1}}_{n-1 \text{ low signals}} \times \underbrace{\frac{s_H}{1-s_H}}_{\text{a single high signal}} \geq \frac{c}{1-c}
\end{aligned}
$$

When the LHS above exceeds the RHS, *adverse selection poses no threat* to evaluators: even after $n-1$ low signals, a high signal suffices to justify an applicant's approval. This threat subsides when evaluators have a lower bar for rejecting the applicant; either due to a favourable prior belief ρ about his quality, or a low approval cost c. Having less evaluators helps too, capping the number of low signals the applicant can accumulate. But fundamentally, whether adverse selection poses a threat depends on the strength of good and bad news. As bad news get stronger against good news, the applicant's $n-1$ low signals increasingly dominate over the single high signal he received. Once the strength of bad news exceeds a threshold, this high signal no longer vindicates the applicant. I denote the normalised belief which marks this threshold as s_L^{as}.

Definition 2. For a binary experiment $\mathcal{E}$ with given strength of good news s_H, s_L^{as} is the *strongest* bad news under which *adverse selection poses no threat:*

$$\frac{\rho}{1-\rho} \times \left(\frac{s_L^{\text{as}}}{1-s_L^{\text{as}}}\right)^{n-1} \times \frac{s_H}{1-s_H} = \frac{c}{1-c}$$

This threshold is intimately related to the cutoff to which Theorem 1 points. Indeed, unless they already approve *every* applicant, stronger bad news always pushes evaluators to approve their *marginal rejects*. Whenever adverse selection poses a threat, these approvals hurt their payoffs. The example in the beginning of this section illustrates this. There, we have no equilibria where *every* applicant is approved; a low signal always warrants a rejection as evaluators' prior belief ρ is already below their approval cost c. Adverse selection poses a threat beyond the threshold $s_L^{\text{as}} = 0.2$; precisely the strength of bad news in their original experiment $\mathcal{E}$. Consequently, switching to $\mathcal{E}^b$ hurts evaluators' payoffs with each marginal reject it pushes them to approve.

The exception is equilibria where bad news are too weak to convince evaluators oblivious about the applicant's past to reject him. Until the strength of bad news exceeds a threshold s_L^{mute} I introduce below, its weakness causes evaluators to approve *every* applicant in the least selective equilibrium. This threshold is in general higher for the most selective equilibrium, where adverse selection might push evaluators to reject low signals earlier.

Definition 3. For a binary experiment $\mathcal{E}$, s_L^{mute} is the *strongest* bad news can get before all equilibria where every applicant is approved vanish:

$$\frac{\rho}{1-\rho} \times \frac{s_L^{\text{mute}}}{1-s_L^{\text{mute}}} = \frac{c}{1-c}$$

Pushing evaluators away from such equilibria might, in general, require strengthening bad news beyond the threshold s_L^{as}. Once they do so however, further strengthening bad news is followed with the approval of marginal rejects, as before.

Proposition 3. Suppose the experiment $\mathcal{E}$ is binary with outcomes inducing the normalised beliefs $s_L \in [0, 0.5]$ and $s_H \in [0.5, 1]$. Stronger bad news decrease evaluators' payoffs when:

i s_L is below the cutoff $\check{s}_L := \min\{s_L^{\text{mute}}, s_L^{as}\}$ in the least selective equilibrium.

ii s_L is below a cutoff $\hat{s_L} \geq \check{s_L}$ in the most selective equilibrium.

Finally, let us turn to the case of *stronger good news.* Now, say the possible outcomes of experiment $\mathcal{E}'$ induce the normalised beliefs:

$$s'_L = \frac{p'_H(s'_L)}{p'_H(s'_L) + p'_L(s'_L)} = s_L \qquad\qquad s'_H = \frac{p'_H(s'_H)}{p'_H(s'_H) + p'_L(s'_H)} = s_H + \delta$$

for some small $\delta > 0$. This new experiment thus offers *marginally stronger good news* than the original experiment $\mathcal{E}$, but the same strength of bad news as the latter. As before, let evaluators approve upon the high outcome and reject upon the low in either experiment. Which applicants' outcome does evaluators' switch to this experiment with stronger news affect, then?

As before, we can interpret the additional information experiment $\mathcal{E}'$ provides as an auxiliary signal evaluators observe following their initial experiment $\mathcal{E}$. This time, they observe this signal $\hat{S}$ only if the outcome of their initial experiment is high, $s = s_H$. This auxiliary signal $\hat{S}$, binary as before, has a distribution:

$$\mathbb{P}\left(\hat{s} = \hat{s_L} \mid \theta = H\right) = \varepsilon \times \frac{s_L}{1 - s_L} \qquad\qquad \mathbb{P}\left(\hat{s} = \hat{s_L} \mid \theta = L\right) = \varepsilon \times \frac{s_H}{1 - s_H}$$

and is, conditional on the applicant's quality θ, independent both from an evaluator's original signal S and anything other evaluators observe. Observing the low auxiliary signal, $\hat{s} = \hat{s_L}$, is already equivalent to observing the low outcome $s' = s'_L$ from the new experiment $\mathcal{E}'$:

$$\frac{\mathbb{P}\left(s = s_H, \hat{s} = \hat{s_L} \mid \theta = H\right)}{\mathbb{P}\left(s = s_H, \hat{s} = \hat{s_L} \mid \theta = L\right)} = \frac{s_H}{1 - s_H} \times \frac{\frac{s_L}{1-s_L}}{\frac{s_H}{1-s_H}} = \frac{s_L}{1 - s_L}$$

We choose ε so that observing the high auxiliary signal $\hat{s} = \hat{s_H}$ is equivalent to observing the high outcome from the new experiment $\mathcal{E}'$:

$$\begin{aligned}\frac{\mathbb{P}\left(s = s_H, \hat{s} = \hat{s_H} \mid \theta = H\right)}{\mathbb{P}\left(s = s_H, \hat{s} = \hat{s_H} \mid \theta = L\right)} &= \frac{\mathbb{P}\left(s' = s'_H \mid \theta = H\right)}{\mathbb{P}\left(s' = s'_H \mid \theta = L\right)} \\ \iff \frac{s_H}{1 - s_H} \times \frac{1 - \varepsilon \times \frac{s_L}{1-s_L}}{1 - \varepsilon \times \frac{s_H}{1-s_H}} &= \frac{s_H + \delta}{1 - s_H - \delta}\end{aligned}$$

as before, the value of ε which ensures this is a continuous and increasing function of δ. I illustrate this construction in the right panel of Figure 3.

This time, the auxiliary signal serves as a "re-evaluation" of an evaluator's initial approval decisions. Her initial rejections are final, but this re-evaluation might overturn an initial approval verdict. This might lose an applicant *any* approval he would otherwise secure, leaving him with a rejection from every evaluator.

Inferring the signals evaluators observed for these *marginal admits* they now reject is easier. This marginal admit is turned down by every evaluator once they observe the auxiliary signals. Thus, every evaluator must have seen a low signal, either $s = s_L$ or $\hat{s} = \hat{s_L}$. Evaluators' payoffs are guaranteed to improve when they reject such an applicant, unless bad news from the low signal are too weak to incentivise the rejection of any applicant. To prove this second part of Theorem 1, I show that strengthening good news pushes evaluators towards such equilibria. Once there, evaluators always benefit from the marginal admits it pushes out.

3.2 Blackwell Improvements of Experiments with Finite Outcomes

In the previous section, I characterised the effect of moving from one binary experiment, $\mathcal{E}$, to a more informative one, $\mathcal{E}'$, on evaluators' payoffs. I showed that the *direction* in which information improves, specficially whether *good news* or *bad news* get stronger, determines this effect. *Stronger good news* always benefits evaluators. It drives them to reject their marginal admits, re-assessed negatively by every evaluator upon richer information. In contrast, *stronger bad news* eventually hurt evaluators. It pushes them to approve their marginal rejects, re-assessed positively by some evaluators upon richer information. These applicants are adversely selected; their positive re-assessments come against the backdrop of other evaluators' negative verdicts. Once bad news are *too strong* against good news, adverse selection overpowers the positive re-assessment that led to their approval, and leaves evaluators worse off.

i don't feel i have clarified that these are the only things you can do to a binary experiment.

Binary experiments afford significant tractability. However, in many settings of interest, evaluators have richer sources of information. Traders of financial assets, for instance, might get recommendations of varying levels of strength such as "Strong Sell", "Sell", "Buy" and "Strong Buy". Likewise, a bank's credit scoring algorithm might output varying probabilities for a loan seeker's default rather than a simple "Good" score and a "Bad" one. With this motivation, here I investigate how the insight in Theorem 1 extends to improvements of *arbitrary* experiments.

shd i mention the restriction to finite outcomes? i shd discuss why this restriction sometime.

Two complications challenge this exercise. Both stem from the fact that normalised beliefs are, in general, ill suited to characterise experiments. The distribution of their outcomes, described by $2m$ unknowns $\{p_L(s), p_H(s)\}_{s\in\mathbf{S}}$, cannot be determined by the $m+2$ linear equalities their normalised beliefs supply:

$$s = \frac{p_H(s)}{p_H(s) + p_L(s)} \quad \text{for } s \in \mathbf{S}$$

$$\sum_{s\in\mathbf{S}} p_\theta(s) = 1 \quad \text{for } \theta \in \{L, H\}$$

A binary experiment is the exception. Its normalised beliefs, which I dubbed s_H and s_L, uniquely determine its outcomes' distributions. Consequently, they describe its informational content. I reinterpreted these normalised beliefs as the strength of good news and bad news an experiment carries, respectively. A more informative binary experiment is simply one with stronger good news and stronger bad news; higher s_H and lower s_L.

I exploited this exception in two ways. First, I characterised Blackwell improvements of a binary experiment through a dichotomy: those which *strengthen good news*, and others which *strengthen bad news*. Any improvement of a binary experiment combines these two forces. This dichotomy was the key to understanding the different Blackwell improvements' effect on applicants, and thus on evaluators' payoffs. How this categorisation extends beyond binary experiments is not immediately clear. Besides being insufficient to characterise improvements in information, normalised beliefs do not have natural interpretations as "good" and "bad" news in general. As the example immediately preceding Proposition 2 showed, evaluators' interpretation of a signal depends very much on the equilibrium they play.

Second, through marginal adjustments to these normalised beliefs, I constructed "small" improvements in evaluators' information. These small improvements in information formed the building blocks of "larger" Blackwell improvements. Most importantly, they helped me uncover how improving an experiment in either direction affects applicants, and ultimately evaluators. Beyond binary experiments, we cannot equate "small" movements in an experiment's normalised beliefs with "small" improvements in the information it carries.

should i replace the phrase "marginal improvement" i used earlier?

The solution to these challenges come from the *auxiliary signals* we constructed in the previous section. Earlier, we used them to replicate and visualise Blackwell improvements of evaluators' experiments, and pin down the applicants they affect. *Local mean preserving spreads* formalise and generalise the auxiliary signals we constructed to the setting of arbitrary experiments. They prove to be the key tool for the remainder of our analysis.

Definition 4 (Local Mean Preserving Spread)**.** Take two experiments $\mathcal{E} = (\mathbf{S}, p_L, p_H)$ and $\mathcal{E}' = (\mathbf{S}', p'_L, p'_H)$ and let $s_1 < s_2 < ... < s_M$ be the normalised beliefs their joint outcome set $\mathbf{S} \cup \mathbf{S}'$ induces. Define the probability distributions p and p' over these outcomes as:

abusing language: sets don't induce normalised beliefs.

$$p(s) := \frac{p_H(s) + p_L(s)}{2} \qquad p'(s) := \frac{p'_H(s) + p'_L(s)}{2} \qquad \text{for all } s \in \{s_1, s_2, ..., s_M\}$$

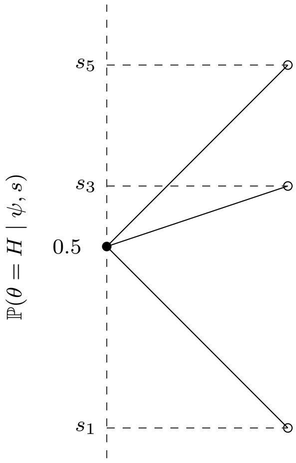
Experiment $\mathcal{E}'$ differs from $\mathcal{E}$ by a *local mean preserving spread* at s_j if for some $j \in \{2, ..., M-1\}$:

$$p'(s_{j-1}) \geq p(s_{j-1}) \qquad 0 = p'(s_j) \leq p(s_j) \qquad p'(s_{j+1}) \geq p(s_{j+1})$$

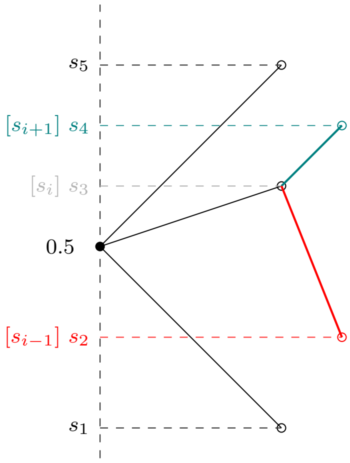
$$p'(s_k) = p(s_k) \quad \text{for all } j \notin \{j-1, j, j+1\}$$

$$\sum_{j=1}^{M} s_j \times p'(s_j) = \sum_{j=1}^{M} s_j \times p(s_j)$$

the original definition has F not $\mathcal{E}$, is that a problem?



(a) Before Local Mean Preserving Spread



(b) After Local Mean Preserving Spread

Figure 4: A Local Mean Preserving Spread

Much like an ordinary mean preserving spread (Rothschild and Stiglitz, 1970[5]), a *local mean preserving spread* distributes probability away from an *origin* point to two *destination* points, one above and one below it. It does so while preserving the mean of the original distribution. Crucially however, a mean preserving spread is *local* if and only if the destination points are the immediate neighbours of the origin point[6]. In other words, neither the original nor the resulting distribution assign positive probability to any other point between the origin and the two destination points[7].

The auxiliary signals I constructed in the previous section generate such local mean preserving spreads (hereafter, just *local spread*). To strengthen bad news, for instance, the auxiliary signal spreads all the probability mass the experiment $\mathcal{E}$ assigns to the origin point s_L to the neighbouring destination points s_H and s'_L, the former above, and the latter below it.

Local spreads are simple Blackwell improvements of an experiment. Like auxiliary signals, we can interpret them as an additional binary experiment an evaluator observes after a particular outcome of her original experiment. Despite their simplicity, local spreads are powerful enough to be pieced together into *any* Blackwell improvement of an experiment. I restate this result in Remark 1. Note that Remark 1 is a slight refinement of Rotschild and Stiglitz's classic result (1970) for experiments with finitely many outcomes.

Remark 1. [Müller and Stoyan, 2002, Theorem 1.5.29] An experiment $\mathcal{E}'$ is Blackwell more informative than another, $\mathcal{E}$, if and only if there is a finite sequence of experiments $\mathcal{E}_1, \mathcal{E}_2, ..., \mathcal{E}_k$

[5]Rothschild and Stiglitz, 1970 describe mean preserving spreads through *four* points in the support of the distribution. Here, I describe them through *three*. This is without loss of generality. In fact, mean preserving spreads were first characterised by Muirhead, 1900 in the context of majorisation, with *three* points. Rasmusen and Petrakis, 1992 show formally that these the three or four point characterisations of MPS are in fact equivalent.

[6]The reader will notice that this statement is ill-defined unless the signal structure is discrete. To the best of the author's knowledge, no counterpart for *local mean preserving spreads* exist for, say, atomless signal structures.

[7]The attentive reader will also realise that this definition also requires that *all* probability mass be spread away from the origin point. This difference is insignificant in our current setting.

such that $\mathcal{E}_1 = \mathcal{E}$, $\mathcal{E}_k = \mathcal{E}'$, and $\mathcal{E}_{i+1}$ differs from $\mathcal{E}_i$ by a local mean preserving spread.

Again, the power of local spreads lie in the way we reinterpret them. A local spread acts as a re-evaluation of an applicant for whom an evaluator initially observes the signal it spreads. Upon this re-evaluation, the evaluator might revise the verdict she would have reached after that signal. This jeopardises an applicant she would approve upon that signal, while offering one she would reject a second chance. In the case of binary experiments, we uncovered that stronger good news have the former influence, and stronger bad news the latter. These different influences drive their different consequences for evaluators' payoffs. The consequence of a local spread, too, depends on which of these influences it exerts on evaluators' verdicts.

Theorem 2. Say the experiment $\mathcal{E}'$ differs from $\mathcal{E}$ by a local mean preserving spread at s_j. Evaluators' payoffs under the most (least) selective equilibrium of $\mathcal{E}'$:

1. *weakly exceed* those under $\mathcal{E}$ if s_j leads to approvals under $\mathcal{E}$; i.e. $\hat{\sigma}_{\mathcal{E}}(s_j) = 1$ ($\check{\sigma}_{\mathcal{E}}(s_j) = 1$)
2. *fall weakly below* those under $\mathcal{E}$ if:

   i if s_j leads to rejections under $\mathcal{E}$; i.e. $\hat{\sigma}_{\mathcal{E}}(s_j) = 1$ ($\check{\sigma}_{\mathcal{E}}(s_j) = 1$), and

   ii the following condition holds:

$$\frac{\rho}{1-\rho} \times \left(\frac{r_H(\sigma;\mathcal{E})}{r_L(\sigma;\mathcal{E})}\right)^{n-1} \times \frac{s_{j+1}}{1-s_{j+1}} \leq \frac{c}{1-c}$$

Evaluators' payoffs always increase when a signal they interpret as *good news* – and approve their applicants upon – is locally spread. This new information might lead all of them to negatively re-assess an applicant who, previously, some would have approved. They benefit from his rejection. In contrast, locally spreading a bad news signal, if the bad news it carries is already too strong, hurts evaluators. Such spreads push evaluators to re-assess applicants they would initially have rejected. This offers applicants who would otherwise be rejected by every evaluator a second chance. Any rejection he overturns following an evaluator's favourable re-assessment comes at the backdrop of others' negative assessments, potentially reinforced following this local spread. When bad news are too strong, the latter negative assessments outweigh the former favourable one. His approval thus hurts evaluators' payoffs.

is the leading discussion clear enough?

check that last condition involves s_{j+1}

Note the similarity between the condition in Theorem 2 determining when stronger bad news hurt evaluators, and the *adverse selection poses no threat* condition we laid out in the previous section. Both conditions hinge on whether the *last* evaluator to see the applicant benefits from approving him, given her signal. Evaluators are tightly bound by adverse selection in the benefit they can extract from better information: locally spreading a bad news signal harms evaluators *whenever* the local spread in concern presents an adverse selection threat.

Theorem 2 relies on the analyst's knowledge of the equilibrium in concern, unlike Theorem 1. The analyst must both know how evaluators interpret the signal that is locally spread, and the probability with which each evaluator rejects a *High* and *Low* quality applicant. In practice, she might wish to remain agnostic about these details pertaining to the form equilibrium takes. To alleviate her concern, I offer a stronger sufficient condition for when a local spread harms evaluators in Proposition 4. This sufficient condition reduces the analyst's dependance on her knowledge of the equilibrium being played.

Proposition 4. Let the experiment $\mathcal{E}'$ differ from $\mathcal{E}$ by a local mean preserving spread at s_j. Furthermore, let σ' and σ be the most selective equilibrium strategies under these two experiments. Evaluators' payoffs are lower with the strategies σ' if:

1. the signal s_j leads to rejections under σ; $\sigma(s_j) = 0$, and
2. the following condition holds:

$$\frac{\rho}{1-\rho} \times \left(\frac{s_j}{1-s_j}\right)^{n-1} \times \frac{s_{j+1}}{1-s_{j+1}} \leq \frac{c}{1-c}$$

The second condition in Proposition 4 strengthens the adverse selection condition Theorem 2 offers. The first condition still relies on the analyst knowing that evaluators interpret s_j as bad news. However, this condition can be strengthened yet further; the condition $s_j < s_L^{\text{mute}}$ is sufficient for s_j to be interpreted as bad news in any equilibrium. When $s_j < s_L^{\text{mute}}$, the signal s_j warrants an applicant's rejection even absent adverse selection.

4 Designing Evaluators' Information

In Section 3, I showed that despite letting them tailor their decision better to the applicant's quality, more information might hurt evaluators' equilibrium payoffs. This owes to an *adverse selection externality* inherent to decentralised evaluations: which applicant one evaluator rejects determines who the other is wary of receiving. With better informed peers, each evaluator might end up expecting a *worse* applicant; wary that some of those peers likely rejected him earlier. Eventually this might leave all evaluators worse off *despite* being better equipped to judge the applicant's quality.

Somewhat paradoxically then, evaluators could benefit from committing to ignore some information about their applicant's quality. Individually, they might be unable to do so. However, a regulator interested in the quality of evaluators' selection might have the power to prescribe both *what* information they must use, and *how* they must use it[8]. In this Section, I study how

[8] **Add to introduction:** Financial regulation in the aftermath of the 2008 crisis provides an example. To improve their lending decisions, regulators sought precisely to prescribe what information lenders can use, and

this regulator can *best* coarsen evaluators' information about the applicant's quality, with the aim of maximising their equilibrium payoffs. Becoming less selective never benefits evaluators (Proposition 2), so in designing their information, the regulator seeks to boost their selectivity. Her approach is aggressive: she tries to steer evaluators away from *any* applicant they might regret approving upon learning his rejection history.

does this feel complete? or awkward end?

Formally, I consider a regulator who garbles evaluators' experiment $\mathcal{E}$ at the outset of the game. She can choose any garbling $S^G : \mathbf{S} \to \Delta\left(\mathbf{S}^G\right)$ which maps the original outcome $s \in \mathbf{S}$ of this experiment to a distribution over $\mathbf{S}^G$, an arbitrary finite set of garbled signals. Previously, I denoted the distribution of experiment $\mathcal{E}$'s outcomes conditional on the applicant's quality as p_θ. Following that notation, I denote the distribution of the garbled signal conditional on experiment $\mathcal{E}$'s outcome $s \in \mathbf{S}$ as $p_s^G := S^G(s)$. Likewise, I denote a representative element of $\mathbf{S}^G$ as s^G.

Once the regulator chooses the garbling S^G, the game proceeds as before. The applicant sequentially visits the n evaluators in the manner described earlier. The evaluator he visits no longer observes the outcome of $\mathcal{E}$, but instead the garbled signal s^G. She then decides whether to approve or reject the applicant. The game ends either when one evaluator approves the applicant, or they all reject him.

An *equilibrium*, as before, is a pair $\left(\sigma^G, \psi^G\right)$ such that (i) the strategy $\sigma^G : \mathbf{S}^G \to [0,1]$ is *optimal* given the interim belief ψ^G, and (ii) the interim belief ψ^G is *consistent* with the strategies σ^G. Throughout this section, I assume that evaluators always play the *most selective equilibrium* that the regulator's chosen garbling supports[9].

The regulator whishes to maximise evaluators' (most selective) equilibrium payoffs through her chosen garbling. In this section, I identify the *regulator-preferred* garbling, which achieves that objective.

I begin by introducing a special class of garblings which are crucial to this exercise.

Definition 5. A garbling S^G of $\mathcal{E}$ is *monotone binary* if:

1. The set $\mathbf{S}^G$ is binary, with elements labelled s_L^G and s_H^G without loss.

2. Either there is a threshold outcome $s^* \in \mathbf{S}$ such that:

$$p_s^G\left(s_H^G\right) = \begin{cases} 0 & s < s^* \\ \in (0,1] & s = s^* \\ 1 & s > s^* \end{cases}$$

how, when assessing borrowers' creditworthiness [details about AtR and Basel III]. In protest, the frustrated lenders highlighted the paradoxical nature of curbing the information required to assess borrowers to improve those same assessments. Discuss the banking chairman's speech. We saw however, that this paradox is only apparent: constraining evaluators' information can improve the quality of their eventual selection by alleviating the adverse selection burden they bear.

[9]Later, I discuss some difficulties that arise when evaluators instead choose the least selective equilibrium.

or $p_s^G(s_H^G) = 0$ for all outcomes $s \in \mathbf{S}$.

Monotone binary garblings are perhaps the simplest that come to mind. Rather than directly observing the outcome of her original experiment $\mathcal{E}$, the evaluator receives either a "high" or a "low" signal from the planner's garbling, which I denote as s_H^G and s_L^G. Outcomes which exceed the threshold $s^* \in \mathbf{S}$ trigger the "high" signal, whereas those below the threshold trigger the "low" one. The threshold outcome s^* is the lowest that can trigger the high signal. Moreover, it is the unique outcome which can *also* trigger the low signal with some probability.

The high signal from such a garbling serves as an "approval recommendation" for evaluators, while the low signal serves as a "rejection recommendation". Of course, the evaluators will judge whether it is *optimal* for them to follow these recommendations. This depends both on what each recommendation tells evaluators about the applicant's quality, and what interim belief following them induces. When evaluators judge following them optimal, I say the planner's garbling *produces optimal recommendations.*

i didn't want to use IC, to tie it to "optimality" of strategies.

Definition 6. A binary garbling S^G produces *optimal recommendations* if the strategy σ^G which obeys its recommendations, defined as:

$$\sigma^G(s^G) := \begin{cases} 0 & s^G = s_L^G \\ 1 & s^G = s_H^G \end{cases}$$

is an equilibrium strategy under S^G.

Despite their simplicity, I establish in Lemma 3 that the regulator need not look beyond monotone binary garblings.

Lemma 3. Whenever it exists, the regulator-preferred garbling is *monotone binary* and *produces optimal recommendations.*

That the regulator can limit herself to binary garblings follows from a fundamental principle in information design. An evaluator ultimately distils the information relayed by the garbled signal into which action she ought to take. The regulator can distil that information herself, sending her only a simple "approve" or "reject" recommendation instead[10]. *Monotone* binary garblings recommend a rejection upon "low" outcomes, and an approval upon "high" outcomes. In doing so, they ensure a better alignment between evaluators' actions with the applicant's quality, in line with the regulator's objective. However, the regulator must still ensure that those recommendations remain optimal. Lemma 3's novelty lies there: I show that she need not

[10]This elementary principle relies crucially on the evaluators coordinating on the most selective equilibrium (equivalently, the regulator's preferred one). Coarsening evaluators' experiment might also *create* new equilibrium outcomes that the regulator might dislike. The regulator might then want to supply evaluators with information beyond which actions she recommends, so that her preferred outcome is the *unique* equilibrium.

depart from monotonicity to do so. For any garbling, there is a monotone binary alternative with optimal recommendations that raises evaluators' equilibrium payoffs.

Much like monotone strategies, monotone binary garblings group the outcomes of $\mathcal{E}$ to an "approval" and "rejection" region. Thus, "selectivity" offers a natural comparison among them. I call a monotone binary garbling *more selective* when it is less likely to send an approval recommendation to evaluators.

There are 8 defns so far. 4 of them here. problem-atic? how to solve?

Definition 7. Where S^G and $S^{G'}$ are monotone binary garblings of $\mathcal{E}$, $S^{G'}$ is *more selective than* S^G if $p_s^{G'}(s_H^{G'}) \leq p_s(s_H^G)$ for any $s \in \mathbf{S}$.

Evaluators' equilibrium strategies, like the planner's preferred garbling, are monotone. They never benefit from moving to a less selective strategy, as Proposition 2 established. But the regulator might rather they approve *more selectively.* As Theorems 1 and 2 foreshadowed, she may want evaluators to "ignore" information that makes them less selective to that end; particularly when an evaluator might regret approving her applicant were she to learn that she was the *only* one not to reject him. I say a monotone binary garbling has regret-free approvals if it *never* recommends an approval upon an outcome $s \in \mathbf{S}$ that might cause such regret.

need a new term here. congruent with "A.S. poses no threat". Dan disliked latter bc "threat" sounds like repeated games.

Definition 8. The monotone binary garbling S^G is said to have *regret-free approvals* if it either:

i Recommends no approvals; i.e. $p_s(s_H^G) = 0$ for all $s \in \mathbf{S}$, or

ii The following condition holds:

$$\frac{p_H(s^*)}{p_L(s^*)} \times \left(\frac{\sum_{s \in \mathbf{S}} p_H(s) \times p_s^G\left(s_L^G\right)}{\sum_{s \in \mathbf{S}} p_L(s) \times p_s^G\left(s_L^G\right)} \right)^{n-1} \geq \frac{c}{1-c}$$

Otherwise, it is said to have *regret-prone approvals.*

what should be the name?

Regret-free approvals is a stringent condition. It demands that approving the applicant upon the high signal s_H^G be optimal for an evaluator regardless of (i) how many rejections the applicant had until he visited her, and (ii) the outcome $s \in \mathbf{S}$ which led to the approval recommendation. Nonetheless, Proposition 5 establishes that the regulator seeks precisely this condition when she is not constrained by the optimality of her recommendations.

Proposition 5. Let S^{G*} be the least selective monotone binary garbling with regret-free approvals, and σ^{G*} be the strategy that obeys its recommendations. Evaluators' payoffs under the strategy σ^{G*} and garbling S^{G*} exceed those under any other strategy and garbling pair.

Short paragraph about why, and why this is interesting.

An evaluator always finds a regret-free approval recommendation optimal. More so, she would find it optimal even if she believed the worse about *when* she received the applicant – after all her peers rejected him – and *why* she was recommended to approve him – because the threshold outcome $s^* \in \mathbf{S}$ materialised. However, evaluators might find S^{G*} *too selective*, preferring to approve their applicants even after rejection recommendations. The regulator must then tailor S^{G*} towards a second-best garbling with optimal recommendations.

Proposition 6. The regulator-preferred garbling is S^{G*} if its recommendations are optimal. Otherwise, it is either:

- the least selective monotone binary garbling with regret-free approvals, or
- the most selective monotone binary garbling with regret-prone approvals

among those with optimal recommendations.

Rather counter-intuitively, the regulator might need to tailor her first-best garbling, S^{G*}, towards an *even more selective* garbling. Such a garbling raises the approval bar even further, but might justify this higher bar by inducing a worse interim belief. For optimal recommendations, the regulator must either recommend rejecting some applicants who would merit an approval, or recommend approving others who would not. Which is the lesser sacrifice is *a priori* unclear.

The regulator is not always constrained by evaluators' optimality constraints, however. An important case is when evaluators find rejecting an applicant optimal absent any information; that is, when their approval cost c weakly exceeds their prior ρ. Since evaluators' interim beliefs are always below their prior, *any* rejection recommendation from a monotone binary garbling – including S^{G*} – is guaranteed to be optimal.

Corollary 7. When evaluators' approval cost c is weakly above their prior ρ, S^{G*} is the regulator-preferred garbling.

should i add examples from binary?

5 Information with Arbitrary History Signals

I am downgrading this in priority, might even be absent. This is because the paper is now called "decentralised evaluations", so i can talk about centralising them in another paper. Nonetheless, there are already results I proved about this. I might add them as an extension. What do you think?

This Section is about the generalisation to arbitrary history signals that I wish to make. I conjecture that Theorem 1 is going to generalise in some form to *any* history signal. I do not wish to get a fully general result in the spirit of Theorem 2; i.e. will restrict myself to binary in this section. This is because with partially (or fully) observed past decisions, Blackwell improvements might behave weirdly simply because decisions are censored data à la classic social learning anyway.

6 Take-It-Or-Leave-It Price Offers

Nothing changes when evaluators offer take it or leave it prices to applicants. Diamond's paradox kicks in, every evaluator offers max acceptable price.

7 Competing in Application Costs

Evaluators post application costs (potentially negative). Applicant applies from lowest to highest cost. Turns the game into an all-pay auction to evade adverse selection, à la Broecker, 1990. Equilibrium strategies of application costs are mixed with no atoms, so ex-post order of applications are perfectly known.

8 Proof Appendix

8.1 Useful Definitions and Notation

In what follows, I occasionally operate with the likelihood ratios of beliefs for convenience. The reader can easily verify the identities:

$$\frac{\psi}{1-\psi} = \frac{\rho}{1-\rho} \times \frac{\nu_H(\sigma;\mathcal{E})}{\nu_L(\sigma;\mathcal{E})} \qquad \frac{\mathbb{P}_\psi(\theta = H \mid s_i)}{1-\mathbb{P}_\psi(\theta = H \mid s_i)} = \frac{\psi}{1-\psi} \times \frac{s_i}{1-s_i}$$

Through similar reasoning, the reader can verify that it is optimal to approve the applicant when:

$$\frac{\mathbb{P}_\psi(\theta = H \mid s_i)}{1-\mathbb{P}_\psi(\theta = H \mid s_i)} > \frac{c}{1-c}$$

Some strategies require evaluators to randomise when approving their applicant upon observing a particular signal realisation. To facilitate the technical discussion, I assume that each evaluator observes the realisation of a *tie-breaking signal* $u \sim U[0,1]$ alongside the outcome of her experiment. This signal is not informative about the applicant's quality: it is distributed independently from it conditional on the experiment's outcome. I denote the outcome of evaluator i's experiment as s^i and her tie-breaking signal as u^i. Without loss, evaluator i approves the applicant if and only if $\sigma(s^i) \leq u^i$; where σ is her strategy. I call the pair (s^i, u^i) the *score* evaluator i observes for the applicant.

Definition 9. The tuple $Z^i = (s^i, u^i)$, where $u^i \overset{IID}{\sim} U[0,1]$ is the *score* evaluator i observes for the applicant. The applicant's *score profile* $\mathbf{z}$ is the set of scores each evaluator would observe if he were to visit them all; $\mathbf{z} = \{(s^i, u^i)\}_{i=1}^n$. Analogously, the applicant's *signal profile* $\mathbf{s} = \{s^i\}_{i=1}^n$ is the set of outcomes of each evaluator's experiment.

Some proofs in Section 8.3 require comparing interim beliefs across pairs of strategies and experiments; $(\sigma, \mathcal{E})$. For convenience, I define the mapping from such a pair to the interim belief consistent with them as $\Psi(.;\mathcal{E}) : [0,1]^n \to [0,1]$:

i might want to change the capitalised letters if they aren't used meaningfully.

$$\Psi(\sigma;\mathcal{E}) := \frac{\rho \times \nu_H(\sigma;\mathcal{E})}{\rho \times \nu_H(\sigma;\mathcal{E}) + (1-\rho) \times \nu_L(\sigma;\mathcal{E})}$$

Wherever necessary, I treat each strategy $\sigma : \mathbf{S} \to [0,1]$ for an experiment $\mathcal{E}$ as a vector in the compact set $[0,1]^m \subset \mathbb{R}^n$. This is a finite dimensional vector space, so I endow it with the metric induced by the taxicab norm without loss of generality (see Kreyszig, 1978 Theorem 2.4-5):

does the following make sense?

$$||\sigma' - \sigma|| = \sum_{j=1}^{m} |\sigma'(s_j) - \sigma(s_j)| \qquad \text{for any two strategies } \sigma' \text{ and } \sigma$$

Note that the interim belief function $\Psi(.;\mathcal{E})$ is thus a continuous function of evaluators' strategies.

8.2 Omitted Results

Lemma 4. Suppose there is a single evaluator, $n = 1$. Her equilibrium expected payoff under experiment $\mathcal{E}'$ exceed that under $\mathcal{E}$ *regardless of her approval cost* $c \in [0, 1]$ *and prior belief* $\rho \in [0, 1]$ if and only if $\mathcal{E}'$ is (Blackwell) more informative than $\mathcal{E}$.

Proof. The sufficiency part of this Lemma follows from Blackwell's Theorem (Blackwell and Girshick, 1954, Theorem 12.2.2). To show necessity, I fix an arbitrary prior belief ρ for the evaluator.

Let q_j be the posterior belief the evaluator forms about the applicant's quality upon observing the outcome $s_j \in \mathbf{S}$:

$$q_j = \frac{\rho \times s_j}{\rho \times s_j + (1 - \rho) \times (1 - s_j)}$$

Furthermore, let $F(.)$ and $F'(.)$ be the CDFs of posterior beliefs $\mathcal{E}$ and $\mathcal{E}'$ induce, respectively, for this prior belief ρ:

$$F(q) = (1 - \rho) \times \sum_{s \in \mathbf{S}: s \leq q} p_L(s) + \rho \times \sum_{s \in \mathbf{S}: s \leq x} p_H(s)$$
$$F'(q) = (1 - \rho) \times \sum_{s \in \mathbf{S}: s \leq q} p'_L(s) + \rho \times \sum_{s \in \mathbf{S}: s \leq x} p'_H(s)$$

The evaluator's expected payoff under $\mathcal{E}$ is given by:

$$\int_c^1 (q - c) dF(q) = \int_c^1 q dF(q) - c \times (1 - F(c)) = (1 - c) - \int_c^1 F(q) dq$$

An analogous expression gives her expected payoff under $\mathcal{E}'$. Therefore, for her expected payoffs under $\mathcal{E}'$ to exceed those under $\mathcal{E}$ for any $c \in [0, 1]$, we must have:

$$\int_c^1 \big(F(q) - F'(q)\big)\, dq \geq 0$$

which is equivalent to $\mathcal{E}'$ being Blackwell more informative than $\mathcal{E}$[11].

□

8.3 Omitted Proofs

Proposition 1. Where Σ is the set of evaluators' equilibrium strategies:

1. Σ is non-empty and compact.

[11]See Müller and Stoyan, 2002, Theorem 1.5.7. The Blackwell order between signal structures is equivalent to the convex order between the posterior belief distributions they induce; see Gentzkow and Kamenica, 2016.

2. Any equilibrium strategy σ^* is *monotone*: $\sigma^*(s) > 0$ for some $s \in \mathbf{S}$ implies that $\sigma^*(s') = 1$ for every $s' \in \mathbf{S'}$ such that $s' > s$.

3. All equilibria exhibit adverse selection: $\psi^* \leq \rho$ for any interim belief ψ^* consistent with an equilibrium strategy $\sigma^* \in \Sigma$.

Proof. In what follows, I treat each strategy $\sigma : \mathbf{S} \to [0,1]$ as a vector in the compact set $[0,1]^m \subset \mathbb{R}^n$, endowed with the taxicab metric (see the end of Section 8.1). I start by proving that any equilibrium strategy must be monotone and all equilibria exhibit adverse selection. Using these observations, I prove that the set of equilibrium strategies is non-empty and compact.

2. Any equilibrium strategy is monotone.

Any equilibrium strategy σ^* must be optimal against the interim belief ψ^* consistent with it. Whenever $\rho \in (0,1)$, $\psi^* = \Psi(\sigma^*; \mathcal{E}) \in (0,1)$, and so $\mathbb{P}_{\psi^*}(\theta = H \mid s') > \mathbb{P}_{\psi^*}(\theta = H \mid S = s)$ for $s', s \in \mathbf{S}$ such that $s' > s$.

3. All equilibria exhibit adverse selection.

A fortiori, $\Psi(\sigma; \mathcal{E}) \leq \rho$ for any monotone strategy σ. To see this, note that $p_H(.)$ first order stochastically dominates $p_L(.)$ since it likelihood ratio dominates it[12]. Therefore, $\nu_L(\sigma; \mathcal{E}) \geq \nu_H(\sigma; \mathcal{E})$. The result then follows since $\frac{\Psi(\sigma;\mathcal{E})}{1-\Psi(\sigma;\mathcal{E})} = \frac{\rho}{1-\rho} \times \frac{\nu_H(\sigma;\mathcal{E})}{\nu_L(\sigma;\mathcal{E})}$.

1. The set of equilibrium strategies is non-empty and compact.

i The set of equilibrium strategies is non-empty.

Define $\Phi(.) : [0,1]^m \to 2^{[0,1]^m}$ to be the evaluators' *best response correspondence*. $\Phi(.)$ maps any strategy σ to the set of strategies that are optimal against the interim belief $\Psi(\sigma; \mathcal{E})$ it induces:

$$\Phi(\sigma) = \left\{\sigma' \in [0,1]^m : \sigma' \text{ is optimal against } \Psi(\sigma; \mathcal{E})\right\}$$

A strategy σ^* is an equilibrium strategy if and only if it is a fixed point of evaluators' best response correspondence; $\sigma^* \in \Phi(\sigma^*)$. I establish that the correspondence Φ has at least such fixed point through Kakutani's Fixed Point Theorem.

Φ is trivially non-empty; every interim belief has some strategy optimal against it. It is also convex valued; if two distinct approval probabilities are optimal after some outcome $s \in \mathbf{S}$, *any* approval probability is optimal upon that outcome.

The only task that remains is to prove that Φ is upper-semi continuous. For this, take an arbitrary sequence of strategies $\{\sigma_n\}$ such that $\sigma_n \to \sigma_\infty$. Denote the interim beliefs consistent

[12] Theorem 1.C.1 in Shaked and Shanthikumar, 2007.

with these strategies as $\psi_n := \Psi(\sigma_n; \mathcal{E})$. Since $\Psi(.; \mathcal{E})$ is continuous in evaluators' strategies, we also have $\psi_n \to \psi_\infty$ where $\psi_\infty = \Psi(\sigma_\infty; \mathcal{E})$. Now, take a sequence of strategies $\{\sigma_n^*\}$ where $\sigma_n^* \in \Phi(\sigma_n)$. Note that every σ_n^* is monotone since optimality against any interim belief $\psi \in (0,1)$ requires monotonicity. We want to show that Φ is upper semi-continuous; i.e.:

$$\sigma_n^* \to \sigma_\infty^* \implies \sigma_\infty^* \in \Phi(\sigma_\infty)$$

By the Monotone Subsequence Theorem, the sequence $\{\sigma_n^*\}$ has a subsequence $\sigma_{n_k}^* \to \sigma_\infty^*$ of strategies whose norms $||\sigma_{n_k}^*||$ are monotone in their indices n_k. Here I take the case where these norms are increasing, the proof is analogous for the opposite case. Since σ_∞^* is the limit of a subsequence of monotone strategies, it must be a monotone strategy too. Assuming otherwise leads to a contradiction; for any $s, s' \in \mathbf{S}$ such that $s' > s$:

$$\sigma_\infty^*(s) > 0 \ \&\ \sigma_\infty^*(s') < 1 \quad \implies \quad \exists N \in \mathbb{N} \text{ s.t. } \forall\ n_k \geq N \quad \sigma_{n_k}^*(s) > 0 \ \&\ \sigma_{n_k}^*(s') < 1$$

Now let $\bar{s}$ be the highest outcome for which $\sigma_\infty^*(\bar{s}) > 0$. I show that:

- If $\sigma_\infty^*(\bar{s}) \in (0,1)$, then:
$$\frac{\psi_\infty}{1-\psi_\infty} \times \frac{\bar{s}}{1-\bar{s}} = \frac{c}{1-c}$$

- If $\sigma_\infty^*(\bar{s}) = 1$, then:
$$\frac{\psi_\infty}{1-\psi_\infty} \times \frac{s}{1-s} \begin{cases} \leq \frac{c}{1-c} & s < \bar{s} \\ \geq \frac{c}{1-c} & s \geq \bar{s} \end{cases}$$

The first case easily follows by noting that:

$$\sigma_\infty^*(\bar{s}) \in (0,1) \implies \sigma_{n_k}^*(\bar{s}) \in (0,1) \implies \frac{\psi_{n_k}}{1-\psi_{n_k}} \times \frac{\bar{s}}{1-\bar{s}} = \frac{c}{1-c} \implies \frac{\psi_\infty}{1-\psi_\infty} \times \frac{\bar{s}}{1-\bar{s}} = \frac{c}{1-c}$$

for all $n_k \geq N' \in \mathbb{N}$. The second case follows similarly, by noting that $\sigma_\infty^*(\bar{s}) = 1$ and $\sigma_\infty^*(s') = 0$ for all $s' < \bar{s}$ implies $\sigma_{n_k}^*(\bar{s}) > 0$ and $\sigma_{n_k}^*(s') = 0$ for all $n_k \geq N'' \in \mathbb{N}$.

ii The set of equilibrium strategies is compact.

Σ is a subset of $[0,1]^m$ and therefore bounded, hence it suffices to show that is closed. Let $\{\sigma_n^*\}$ be a sequence of equilibrium strategies. Note that this means $\sigma_n^* \in \Phi(\sigma_n^*)$. Since $\Phi(.)$ is upper semi-continuous, $\sigma_n^* \to \sigma_\infty$ implies $\sigma_\infty \in \Phi(\sigma_\infty)$, and therefore an equilibrium strategy itself.

□

Proposition 2. Where σ^* and σ^{**} are two equilibrium strategies such that σ^{**} is more selective than σ^*, evaluators' expected payoffs under σ^{**} exceed those under σ^*; $\Pi(\sigma^{**}; \mathcal{E}) \geq \Pi(\sigma^*; \mathcal{E})$.

Proof. This is an immediate corollary to Lemmas 2 and 5 below; both of independent interest. □

Lemma 5. Take three monotone strategies σ'', σ' and, σ, ordered from the least selective to the most. If $\Pi(\sigma'; \mathcal{E}) \leq \Pi(\sigma; \mathcal{E})$, then $\Pi(\sigma''; \mathcal{E}) \leq \Pi(\sigma'; \mathcal{E})$.

Proof. For the three strategies σ'', σ', and σ, consider three sets $Z, Z', Z'' \subset (S \times [0,1])^n$ where the applicant's score profile $\mathbf{z}$ might lie:

$$\mathbf{z} \in \begin{cases} Z & \text{if } \mathbf{z} \text{ is eventually approved under } \sigma'' \text{ but not } \sigma \\ Z' & \text{if } \mathbf{z} \text{ is eventually approved under } \sigma' \text{ but not } \sigma \\ Z'' & \text{if } \mathbf{z} \text{ is eventually approved under } \sigma'' \text{ but not } \sigma' \end{cases}$$

Notice that $Z' \cap Z'' = \emptyset$ and $Z' \cup Z'' = Z$. We can write the difference between the sum of evaluators' payoffs under different strategies as:

$$\Pi(\sigma'; \mathcal{E}) - \Pi(\sigma; \mathcal{E}) = \mathbb{P}\left(\mathbf{z} \in Z'\right) \times \left[\mathbb{P}\left(\theta = H \mid \mathbf{z} \in Z'\right) - c\right]$$

and:

$$\Pi(\sigma''; \mathcal{E}) - \Pi(\sigma'; \mathcal{E}) = \mathbb{P}\left(\mathbf{z} \in Z''\right) \times \left[\mathbb{P}\left(\theta = H \mid \mathbf{z} \in Z''\right) - c\right]$$

Therefore we want to prove that:

$$\mathbb{P}\left(\theta = H \mid \mathbf{z} \in Z'\right) \leq c \implies \mathbb{P}\left(\theta = H \mid \mathbf{z} \in Z''\right) \leq c$$

Now, note that $\mathbb{P}(\theta = H \mid \mathbf{z} \in Z)$ is a convex combination of $\mathbb{P}(\theta = H \mid \mathbf{z} \in Z')$ and $\mathbb{P}(\theta = H \mid \mathbf{z} \in Z'')$. Furthermore:

$$\mathbb{P}(\theta = H \mid \mathbf{z} \in Z) \geq \mathbb{P}\left(\theta = H \mid \mathbf{z} \in Z \cap Z''\right) = \mathbb{P}\left(\theta = H \mid \mathbf{z} \in Z''\right)$$

which then implies:

$$\mathbb{P}\left(\theta = H \mid \mathbf{z} \in Z''\right) \leq \mathbb{P}(\theta = H \mid \mathbf{z} \in Z) \leq \mathbb{P}\left(\theta = H \mid \mathbf{z} \in Z'\right) \leq c$$

□

Lemma 2. Let σ^* and σ be two monotone strategies, where σ^* is more selective than σ. If σ^* is an equilibrium strategy, then $\Pi(\sigma^*; \mathcal{E}) \geq \Pi(\sigma; \mathcal{E})$.

Proof. Let $\mathbf{z}$ be the applicant's *score profile*. Take an equilibrium strategy σ^* and a less selective strategy σ such that:

$$\sigma(s) - \sigma^*(s) = \begin{cases} \varepsilon & s = \underline{s} \\ 0 & \text{otherwise} \end{cases}$$

for some $\varepsilon > 0$, where $\underline{s} := \min\{s \in S : \sigma^*(s) < 1\}$. I show that:

$$\lim_{\varepsilon \to 0} \Pi(\sigma; \varepsilon) - \Pi(\sigma^*; \varepsilon) \leq 0$$

By Lemma 5, this establishes the result.

Now, let $Z \subset (S \times [0,1])^n$ be the set of score profiles with which at least one evaluator approves the applicant with σ, but all reject him with σ^*:

$$\mathbf{z} \in Z \iff \begin{array}{rl} \sigma^*(s^i) > u^i & \text{for all } i \in \{1, 2, \ldots, n\}, \\ & \text{and} \\ \sigma(s^i) \leq u^i & \text{for some } i \in \{1, 2, \ldots, n\}. \end{array}$$

Furthermore, for a given score profile $\mathbf{z}$, let $\#$ be the number of evaluators whose observed scores are such that $\sigma(s^i) \geq u^i > \sigma^*(s^i)$. These evaluators would approve the applicant under the strategy σ, but not under σ^*.

An applicant's eventual outcome differs between the strategy profiles σ and σ^* if and only if his score profile $\mathbf{z}$ lies in Z. Furthermore, his eventual outcome can only change from a rejection by all evaluators in σ^* to an approval by some evaluator in σ. Thus:

$$\begin{aligned} \Pi(\sigma; \mathcal{E}) - \Pi(\sigma^*; \mathcal{E}) &= [\mathbb{P}(\theta = H \mid \mathbf{z} \in Z) - c] \times \mathbb{P}(\mathbf{z} \in Z) \\ &\propto \mathbb{P}(\theta = H \mid \mathbf{z} \in Z) - c \end{aligned}$$

Focus therefore, on the probability that $\theta = H$ given the applicant's signal profile lies in Z:

$$\mathbb{P}(\theta = H \mid \mathbf{z} \in Z) = \sum_{i=1}^{n} \mathbb{P}(\theta = H \mid \# = i) \times \frac{\mathbb{P}(\# = i)}{\mathbb{P}(\mathbf{z} \in Z)}$$

Now note:

$$\mathbb{P}(\# = i \mid \theta) = (p_\theta(\underline{s}))^i \times (1 - p_\theta(\underline{s}))^{n-i} \times \varepsilon^i$$

and thus $\mathbb{P}(\# = i) \propto \varepsilon^i$. Since $\mathbb{P}(\mathbf{z} \in A) = \sum_{i=1}^{n} \mathbb{P}(\# = i)$, we have $\lim_{\varepsilon \to 0} \frac{\mathbb{P}(\# = i)}{\mathbb{P}(\mathbf{z} \in A)} = 0$ for any $i > 1$. Thus:

$$\lim_{\varepsilon \to 0} \mathbb{P}(\theta = H \mid \mathbf{z} \in A) - \mathbb{P}(\theta = H \mid \# = 1) = 0$$

I conclude the proof by showing that $\mathbb{P}(\theta = H \mid \# = 1) \leq c$ as $\varepsilon \to 0$:

$$\begin{aligned}
\lim_{\varepsilon\to 0} \frac{\mathbb{P}(\theta = H \mid \# = 1)}{\mathbb{P}(\theta = L \mid \# = 1)} &= \lim_{\varepsilon\to 0} \frac{\mathbb{P}(\theta = H)}{\mathbb{P}(\theta = L)} \times \frac{\mathbb{P}(\# = 1 \mid \theta = H)}{\mathbb{P}(\# = 1 \mid \theta = L)} \\
&= \lim_{\varepsilon\to 0} \frac{\mathbb{P}(\theta = H)}{\mathbb{P}(\theta = L)} \times \left(\frac{r_H(\sigma;\mathcal{E})}{r_L(\sigma;\mathcal{E})}\right)^{n-1} \times \frac{p_H(\underline{s})}{p_L(\underline{s})} \\
&= \frac{\mathbb{P}(\theta = H)}{\mathbb{P}(\theta = L)} \times \left(\frac{r_H(\sigma^*;\mathcal{E})}{r_L(\sigma^*;\mathcal{E})}\right)^{n-1} \times \frac{p_H(\underline{s})}{p_L(\underline{s})} \\
&\leq \frac{\psi^*}{1-\psi^*} \times \frac{p_H(\underline{s})}{p_L(\underline{s})} \leq \frac{c}{1-c}
\end{aligned}$$

where $\psi^* = \Psi(\sigma^*;\mathcal{E})$ is the interim belief of the evaluators induced by σ^*. The penultimate inequality holds due to the straightforward fact that:

$$\frac{\psi^*}{1-\psi^*} = \frac{\rho}{1-\rho} \times \frac{1 + r_H^* + ... + (r_H^*)^{n-1}}{1 + r_L^* + ... + (r_L^*)^{n-1}} \leq \frac{\rho}{1-\rho} \times \left(\frac{r_H^*}{r_L^*}\right)^{n-1}$$

where $r_\theta^* := r_\theta(\sigma^*;\mathcal{E})$. The last inequality is due to the fact that $\underline{s} \in S$ is optimally rejected under σ^*.

□

Theorem 1. Let the experiment $\mathcal{E}$ be binary with outcomes inducing the normalised beliefs $s_L \in [0, 0.5]$ and $s_H \in [0.5, 1]$. Increasing s_H weakly increases evaluators' payoffs across the extreme equilibria. In contrast, as s_L decreases, evaluators' payoffs in the most (least) selective equilibrium:

1. weakly improve as long as s_L remains below a cutoff $\hat{s_L}$ ($\check{s_L}$),
2. weakly decrease once s_L falls below this cutoff.

I will use Lemmas 6, 7, and 8 below, possibly of independent interest, to prove Theorem 1. Throughout, I denote the most and least selective equilibrium strategies under the experiment $\mathcal{E}$ as $\hat{\sigma}^*_{\mathcal{E}}$ and $\check{\sigma}^*_{\mathcal{E}}$, respectively. I drop the subscript whenever the experiment in question is obvious.

Lemma 6. Let $\mathcal{E}$ be a binary experiment, with outcomes in $\mathbf{S} = \{s_L, s_H\}$, labelled after the respective normalised beliefs they induce. $\Psi(\sigma;\mathcal{E})$ is:

i strictly increasing in $\sigma(s_L)$, whenever $\sigma(s_H) = 1$, is strictly true here?

ii strictly decreasing in $\sigma(s_H)$ whenever $\sigma(s_L) = 0$.

Proof. Part i:

Let $\sigma(s_L) \in (0,1)$ and $\sigma(s_H) = 1$. The interim belief $\Psi(\sigma; \mathcal{E})$ is then given by:

$$\begin{aligned}
\Psi(\sigma;\mathcal{E}) &= \mathbb{P}\left(\theta = H \mid \text{visit received}\right) \\
&= \sum_{i=0}^{n-1} \mathbb{P}(\text{visited after i}^{\text{th}} \text{ rejection } \mid \text{visit received}) \times \mathbb{E}\left[\theta = H \mid \text{i } s_L \text{ signals}\right] \\
&= \sum_{i=0}^{n-1} \frac{\mathbb{P}(\text{visited after i}^{\text{th}} \text{ rejection })}{\mathbb{P}(\text{visit received})} \times \mathbb{E}\left[\theta = H \mid \text{i } s_L \text{ signals}\right]
\end{aligned}$$

Note that $\mathbb{E}\left[\theta = H \mid \text{i } s_L \text{ signals}\right] < \mathbb{E}\left[\theta = H \mid \text{i+1 } s_L \text{ signals}\right]$; since every s_L signal is further evidence for $\theta = L$. We have:

$$\begin{aligned}
\mathbb{P}(\text{ visited after i}^{\text{th}} \text{ rejection }) =& \mathbb{P}\left(\text{ev. was (i+1)}^{\text{st}} \text{ in order } \mid \text{applicant got i rejections}\right) \\
& \times \mathbb{P}\left(\text{applicant got i rejections}\right) \\
=& \frac{1}{n} \times \mathbb{P}(\text{i } s_L \text{ signals}) \times \left[1 - \sigma(s_L)\right]^i
\end{aligned}$$

The proof is completed by noting that:

$$\frac{\mathbb{P}(\text{ visited after (i+1)}^{\text{st}} \text{ rejection })}{\mathbb{P}(\text{ visited after i}^{\text{th}} \text{ rejection })} = \frac{\mathbb{P}(\text{i+1 } s_L \text{ signals})}{\mathbb{P}(\text{i } s_L \text{ signals})} \times \left[1 - \sigma(s_L)\right]$$

decreases, and thus $\Psi(\sigma; \mathcal{E})$ increases, in $\sigma(s_L)$.

do we need explanation? the distribution over no. low signals FOSD improves.

Part ii:

Now take $\sigma(s_L) = 0$. We then have:

$$r_H(\sigma;\mathcal{E}) = 1 - p_H(s_H)\sigma(s_H) \qquad\qquad r_L(\sigma;\mathcal{E}) = 1 - p_L(s_H)\sigma(s_H)$$

and:

$$\begin{aligned}
\Psi(\sigma;\mathcal{E}) &\propto \frac{1 + r_H + ... + r_H^{n-1}}{1 + r_L + ... + r_L^{n-1}} \\
&= \frac{1 - r_H^n}{1 - r_L^n} \times \frac{1 - r_L}{1 - r_H} = \frac{1 - r_H^n}{1 - r_L^n} \times \frac{p_L(s_H)}{p_H(s_H)} \\
&\propto \frac{1 - r_H^n}{1 - r_L^n} = \frac{1 - \left(1 - p_H(s_H)\sigma(s_H)\right)^n}{1 - \left(1 - p_L(s_H)\sigma(s_H)\right)^n}
\end{aligned}$$

Differentiating the last expression with respect to $\sigma(s_H)$ and rearranging its terms reveals that this derivative is proportional to:

$$\frac{s_H}{1 - s_H} \times \left(\frac{r_H}{r_L}\right)^{n-1} - \frac{1 - (r_H)^n}{1 - (r_L)^n}$$

The positive term is the likelihood ratio of one s_H signal and $n-1$ rejections, and the negative term is the likelihood ratio from *at most* $n-1$ rejections. Since approvals only happen with s_H signals, the negative term strictly exceeds the positive term. This can be verified directly, too:

$$\begin{aligned}\frac{1-(r_H)^n}{1-(r_L)^n} > \frac{s_H}{1-s_H} \times \left(\frac{r_H}{r_L}\right)^{n-1} &\iff \frac{1-(r_H)^n}{1-(r_L)^n} \times \frac{1-r_L}{1-r_H} > \left(\frac{r_H}{r_L}\right)^{n-1} \\ &\iff \frac{1+\ldots+(r_H)^{n-1}}{1+\ldots+(r_L)^{n-1}} > \left(\frac{r_H}{r_L}\right)^{n-1}\end{aligned}$$

The last inequality can be verified easily. Thus, $\Psi(\sigma;\mathcal{E})$ decreases in $\sigma(s_H)$.

□

The Corollary below follows from Lemma 6. Let both $\mathcal{E}'$ and $\mathcal{E}$ are binary experiments, where the former is Blackwell more informative than the latter. If evaluators approve upon the high outcome and reject upon the low in both experiments, the interim belief under $\mathcal{E}'$ is lower.

i am still not sure i convinced the reader of strictness here. check again.

Corollary 8. Let $\mathcal{E}'$ and $\mathcal{E}$ be two binary experiments, where the former is Blackwell more informative than the latter. Let the strategies σ' and σ for these respective experiments be defined as:

$$\sigma'(s') := \begin{cases} 0 & s' = s'_L \\ 1 & s' = s'_H \end{cases} \qquad\qquad \sigma(s) := \begin{cases} 0 & s = s_L \\ 1 & s = s_H \end{cases}$$

Then, $\Psi(\sigma';\mathcal{E}') \leq \Psi(\sigma;\mathcal{E})$.

Proof. Establishing that this holds for a pair $(\mathcal{E}',\mathcal{E})$ for which either (i) $s'_H > s_H$ and $s_L = s'_L$, or (ii) $s'_H = s_H$ and $s_L > s'_L$ suffices. I will only prove the first case, the second is analogous. Below I show that the outcome induced by σ under experiment $\mathcal{E}$ can be replicated by some strategy $\tilde{\sigma}$ under experiment $\mathcal{E}'$, where $\tilde{\sigma}(s_L) > 0$ and $\tilde{\sigma}(s_H) = 1$. Then, the desired conclusion follows from Lemma 6.

Take the pair $(\sigma,\mathcal{E})$. The probabilities that the applicant is rejected or approved upon a visit, conditional on θ, is given by:

$$\frac{\mathbb{P}(\sigma \text{ rejects } \mid \theta = H)}{\mathbb{P}(\sigma \text{ rejects } \mid \theta = L)} = \frac{s_L}{1-s_L} \qquad\qquad \frac{\mathbb{P}(\sigma \text{ approves } \mid \theta = H)}{\mathbb{P}(\sigma \text{ approves } \mid \theta = L)} = \frac{s_H}{1-s_H}$$

For the pair $(\tilde{\sigma},\mathcal{E}')$ where $\tilde{\sigma}(s'_H) = 1$, we have:

$$\frac{\mathbb{P}(\tilde{\sigma} \text{ rejects} \mid \theta = H)}{\mathbb{P}(\tilde{\sigma} \text{ rejects } \mid \theta = L)} = \frac{s_L}{1-s_L} \qquad \frac{\mathbb{P}(\tilde{\sigma} \text{ approves } \mid \theta = H)}{\mathbb{P}(\tilde{\sigma} \text{ approves } \mid \theta = L)} = \frac{p'_H(s_H) + \tilde{\sigma}(s_L)p'_H(s_L)}{p'_L(s_H) + \tilde{\sigma}(s_L)p'_L(s_L)}$$

where $\{p'_L, p'_H\}$ are the distributions pertaining to $\mathcal{E}'$. It is easy to verify that the expression on the right falls from $\frac{s'_H}{1-s'_H}$ to 1 monotonically and continuously as $\tilde{\sigma}(s_L)$ rises from 0 to 1. Thus, there is a unique interior value of $\tilde{\sigma}(s_L)$ that replicates the outcome of $(\sigma;\mathcal{E})$.

□

Lemma 7. Let $\mathcal{E}$ be a binary experiment. The equilibrium approval probability upon the low outcome s_L is either 0 or 1, both in the most and the least selective equilibrium.

Proof. I start by proving this for the least selective equilibrium; i.e. $\check{\sigma}^*(s_L) \in \{0, 1\}$. For s_L^{mute} defined in Definition 3, observe that when $s_L \geq s_L^{\text{mute}}$, $\sigma(s_L) = \sigma(s_H) = 1$ is an equilibrium; so we must have $\check{\sigma}^*(s_L) = 1$. The strategy σ defined by $\sigma(s_H) = \sigma(s_L) = 1$ gives rise to the interim belief $\Psi(\sigma; \mathcal{E}) = \rho$, which in turn renders approving upon the outcome s_L optimal. In turn, if $s_L < s_L^{\text{mute}}$, we must have $\sigma^*(s_L) = 1$ for any equilibrium strategy; since evaluators' interim beliefs always fall below their prior (Proposition 1).

Now consider the most selective equilibrium strategy; $\hat{\sigma}^*$. For contradiction, let $1 > \hat{\sigma}^*(s_L) > 0$ and $\hat{\sigma}^*(s_H) = 1$. Lemma 6 establishes that the interim belief falls as $\sigma(s_L)$ falls; which implies there must be another, more selective equilibrium strategy σ^* such that $\sigma^*(s_L) = 0$ and $\sigma^*(s_H) = 1$.

□

Lemma 7 establishes that when their experiment $\mathcal{E}$ is binary, evaluators *never* mix upon seeing the low outcome $s = s_L$ in extreme equilibria. Following up, Lemma 8 establishes that a more informative binary experiment pushes evaluators to reject upon the low outcome in extreme equilibria.

Lemma 8. Let $\mathcal{E}$ be a binary experiment, with outcomes in $\mathbf{S} = \{s_L, s_H\}$, labelled after the respective normalised beliefs they induce. Evaluators' approval probabilities upon the low outcome $s = s_L$ are given by:

$$\check{\sigma}^*(s_L) = \begin{cases} 1 & s_L \geq s_L^{\text{mute}} \\ 0 & s_L < s_L^{\text{mute}} \end{cases} \qquad \hat{\sigma}^*(s_L) = \begin{cases} 1 & s_L < s_L^{\dagger}(s_H) \\ 0 & s_L \geq s_L^{\dagger}(s_H) \end{cases}$$

where $s_L^{\dagger}(.)$ is an increasing function of s_H, and $s_L^{\dagger}(s_H) \geq s_L^{\text{safe}}$.

Proof. Note that there exists an equilibrium where $\sigma(s_L) = 1$ if and only if:

$$\frac{\rho}{1-\rho} \times \frac{s_L}{1-s_L} \geq \frac{c}{1-c}$$

which, combined with Lemma 7, proves the part of the Lemma for the selective equilibrium.

Now, define the strategies σ_0 as σ_1 as:

$$\sigma_0(s) = \begin{cases} 0 & s = s_L \\ 0 & s = s_H \end{cases} \qquad \sigma_1(s) = \begin{cases} 0 & s = s_L \\ 1 & s = s_H \end{cases}$$

A necessary and sufficient condition for an equilibrium σ^* where $\sigma^*(s_L) = 0$ to exist is:

$$\frac{\Psi(\sigma_1; \mathcal{E})}{1 - \Psi(\sigma_1; \mathcal{E})} \times \frac{s_L}{1 - s_L} \leq \frac{c}{1 - c}$$

Sufficiency follows since either:

$$\frac{\Psi(\sigma_0; \mathcal{E})}{1 - \Psi(\sigma_0; \mathcal{E})} \times \frac{s_H}{1 - s_H} \leq \frac{c}{1 - c}$$

which implies σ_0 is an equilibrium, or there is an equilibrium strategy σ^* such that $\sigma^*(s_L) = 0$ and $\sigma^*(s_H) > 0$ by Lemma 6. The condition is necessary, since any strategy that is less selective than σ_1 induces a higher interim belief, by Lemma 6.

By Corollary 8, whenever this necessary and sufficient condition holds for an experiment $\mathcal{E}$, it also holds for a (Blackwell) more informative experiment $\mathcal{E}'$. Moreover, whenever the low signals are rejected in the least selective equilibrium, they must be in the most selective equilibrium. This concludes the proof.

□

Proof, Theorem 1: By Lemma 8, Blackwell improving evaluators' experiment shifts both their least selective and most selective equilibrium strategies once from approving *every* applicant to rejecting upon the low signal. By Lemma 5, this shift in evaluators' strategy increases evaluators' payoffs.

Let $\{\sigma_\alpha\}_{\alpha \in [0,1]}$ be the family of strategies where evaluators reject upon the low signal:

$$\sigma_\alpha(s) = \begin{cases} 0 & s = s_L \\ 1 & s = s_H \end{cases}$$

By Lemma 6, the interim belief ψ_α that the strategy σ_α induces is strictly decreasing in α. Thus, at most one of these can be an equilibrium strategy for a given experiment. Furthermore, whenever evaluators' expected payoffs from σ_1 is weakly positive, this must be the equilibrium strategy; decreasing α can only make approving upon the high signal *more* profitable. Hence, whenever evaluators reject upon the low signal in equilibrium, their payoffs are given by: $\Pi(\sigma^*; \mathcal{E}) = \max\{0, \Pi(\sigma_1; \mathcal{E})\}$. The Theorem then follows from the Claim below:

Claim. $\max\{0, \Pi(\sigma_1; \mathcal{E})\}$ is:

i weakly increasing in s_H whenever there is some equilibrium strategy σ^* s.t. $\sigma^*(s_L) = 0$.

ii hump-shaped in s_L. As s_L falls, it:

- weakly increases when $s_L \geq s_L^{as}$,
- weakly decreases when $s_L \leq s_L^{as}$

where s_L^{as} is defined in Definition 2.

Proof of the Claim. Part i: Increasing the strength of good news; i.e. s_H.
Let $\mathcal{E}$ and $\mathcal{E}'$ be two binary experiments with outcome sets $\mathbf{S} = \{s_L, s_H\}$ and $\mathbf{S}' = \{s'_L, s'_H\}$; each labelled after the normalised beliefs they induce. The experiment $\mathcal{E}'$ carries *marginally stronger good news* than experiment $\mathcal{E}$:

$$s'_L = s_L \qquad\qquad s'_H = s_H + \delta$$

for some small δ such that $1 - s_H \geq \delta > 0$. I show that $\Pi(\sigma'_1; \mathcal{E}') > \Pi(\sigma_1; \mathcal{E})$; where σ'_1 is defined analogously to σ_1 for experiment $\mathcal{E}'$.

Step 1: Replicating $\mathcal{E}'$ with a signal pair $(s, \hat{s})$.

I am not sure "replication" is a good word to use here.

Rather than observing the outcome of experiment $\mathcal{E}'$, say an evaluator initially observes her original signal s, and then potentially an additional auxilliary signal $\hat{s}$. The first signal she receives, s, records the outcome of $\mathcal{E}$. If the low outcome s_L materialises, the evaluator observes no more information. If, however, the high outcome s_H materialises, she then observes the additional auxiliary signal $\hat{s}$. This auxiliary signal records the outcome of *another* binary experiment, $\hat{\mathcal{E}}$. The outcome of $\hat{\mathcal{E}}$ is independent both from s and anything else any other evaluator observes. Conditional on the applicant's quality θ, the distribution over its outcomes is given by the pmf $p_\theta(.)$:

rethink the hat notation. i am now using it for selectiveness.

$$\hat{p}_H(\hat{s_H}) = 1 - \varepsilon \times \frac{s_L}{1 - s_L} \qquad\qquad \hat{p}_L(\hat{s_H}) = 1 - \varepsilon \times \frac{s_H}{1 - s_H}$$

The evolution of the evaluator's beliefs when she observes this signal pair is determined by the two likelihood ratios:

$$\frac{\mathbb{P}((s,\hat{s}) = (s_H, \hat{s_H}) \mid \theta = H)}{\mathbb{P}((s,\hat{s}) = (s_H, \hat{s_H}) \mid \theta = L)} = \frac{s_H}{1 - s_H} \times \frac{1 - \varepsilon \times \dfrac{s_L}{1 - s_L}}{1 - \varepsilon \times \dfrac{s_H}{1 - s_H}} \tag{8.1}$$

$$\frac{\mathbb{P}((s,\hat{s}) = (s_H, \hat{s_L}) \mid \theta = H)}{\mathbb{P}((s,\hat{s}) = (s_H, \hat{s_L}) \mid \theta = L)} = \frac{s_L}{1 - s_L} \tag{8.2}$$

Note that the likelihood ratio 8.1 increases continuously with ε.

The information from observing the pair $(s, \hat{s})$ as such is equivalent to observing the outcome of experiment $\mathcal{E}'$, when:

$$\frac{s_H}{1 - s_H} \times \frac{1 - \varepsilon \times \dfrac{s_L}{1 - s_L}}{1 - \varepsilon \times \dfrac{s_H}{1 - s_H}} = \frac{s_H + \delta}{1 - (s_H + \delta)} \tag{8.3}$$

for our chosen (δ, ε). I choose ε to satisfy this equality for our δ. As such, ε becomes a continuously increasing function of δ. Furthermore, note that by varying ε between 0 and $\frac{1-s_H}{s_H}$, we can replicate *any* experiment $\mathcal{E}'$ with $s'_L = s_L$ and $1 \geq s'_H \geq s_H$.

Step 2: $\pi(\sigma'_1; \mathcal{E}') \geq \pi(\sigma_1; \mathcal{E})$.

The evaluator who observes the signal pair $(s, \hat{s})$ obtains equivalent information to that from $\mathcal{E}'$. We now must identify the strategy $\tilde{\sigma} : \{s_L, (s_H, \hat{s_H}), (s_H, \hat{s_L})\} \to [0,1]$ for this signal pair that replicates the outcome of the strategy σ'_1 for experiment $\mathcal{E}'$. This strategy is defined as:

$$\tilde{\sigma}(s_H, \hat{s_H}) = 1 \qquad\qquad \tilde{\sigma}(s_L) = \tilde{\sigma}(s_H, \hat{s_L}) = 0$$

and replicates the likelihood ratios of an approval and rejection signal under $\mathcal{E}'$.

Now, fix the applicant's *signal profile* (defined in Section 8.1): the collection of signal draws each evaluator will observe if he visits them all: $\mathbf{s} = \{(s^i, \hat{s}^i)\}_{i=1}^n$. I call an applicant a *marginal admit* if his score profile is such that:

i for at least one $i \in \{1, 2, ..., n\}$, $s^i = s_H$, and

ii for *every* $i \in \{1, 2, ..., n\}$, either $s^i = s_L$, or $\hat{s}^i = \hat{s}_L$.

These marginal admits drive the wedge between evaluators' payoffs under $\mathcal{E}'$ and $\mathcal{E}$: while one of their evaluators approves them under $\mathcal{E}$, they *all* reject him under $\hat{\mathcal{E}}$. So:

$$\Pi(\sigma'_1; \mathcal{E}') - \Pi(\sigma_1; \mathcal{E}) = \mathbb{P}\left(\text{marginal admit}\right) \times \underbrace{\left[c - \mathbb{P}\left(\theta = H \mid \text{marginal admit}\right)\right]}_{(1)}$$

A marginal admit only has signal realisations $(s, \hat{s}) = (s_H, \hat{s_L})$ or $s = s_L$. These carry equivalent information about θ. Thus, the expression (1) above equals:

$$c - \mathbb{P}\left[\theta = H \mid s^1 = ... = s^n = s_L\right]$$

In the relevant region where there is an equilibrium strategy that leads to rejections after the low outcome s_L, the expression above must be weakly positive:

$$\begin{aligned}
c - \mathbb{P}\left[\theta = H \mid s^1 = ... = s^n = s_L\right] &\propto \frac{c}{1-c} \times \frac{\rho}{1-\rho} \times \left(\frac{s_L}{1-s_L}\right)^n \\
&\leq \frac{c}{1-c} - \frac{\rho}{1-\rho} \times \frac{\sum\limits_{k=0}^{n-1} p_H(s_L)^k}{\sum\limits_{k=0}^{n-1} p_L(s_L)^k} \times \frac{s_L}{1-s_L} \\
&= \frac{c}{1-c} - \frac{\Psi(\sigma_1; \mathcal{E})}{1-\Psi(\sigma_1; \mathcal{E})} \times \frac{s_L}{1-s_L} \leq 0
\end{aligned}$$

where the last inequality follows from the necessary and sufficient condition the proof of Lemma 8 introduced for such an equilibrium to exist.

Part ii: Increasing the strength of bad news; i.e. decreasing s_L.

Now, let the experiment $\mathcal{E}'$ carry *marginally stronger bad news* than experiment $\mathcal{E}$ instead; for some arbitrarily small $\delta \in [0, s_L]$:

$$s'_L = s_L - \delta \qquad\qquad s'_H = s_H$$

Where σ'_1 and σ_1 are defined as before, I show that:

i $\Pi(\sigma'_1; \mathcal{E}) - \Pi(\sigma_1; \mathcal{E}) \geq 0$ when $s_L \geq s_L^{as}$, and

ii $\Pi(\sigma'_1; \mathcal{E}) - \Pi(\sigma_1; \mathcal{E}) \leq 0$ when $s_L \leq s_L^{as}$

Step 1: Replicating $\mathcal{E}'$ with a signal pair $(s, \hat{s})$.

As before, let the evaluator potentially observe *two* signals, s and $\hat{s}$. She first observes s, which records the outcome of $\mathcal{E}$. If the high outcome s_H materialises, she receives no further information. If, however, the low outcome s_L materialises, she then observes the additional auxiliary signal $\hat{s}$, which records the outcome of *another* binary experiment, $\hat{\mathcal{E}}$. As before, the outcome of this experiment is independent both from s and anything else any other evaluator observes. Its distribution conditional on the applicant's quality θ is given by the pmf $p_\theta(.)$:

$$\hat{p}_H(\hat{s}_H) = \varepsilon \times \frac{s_H}{1 - s_H} \qquad\qquad \hat{p}_L(\hat{s}_H) = \varepsilon \times \frac{s_L}{1 - s_L}$$

The evolution of the evaluator's beliefs upon seeing the signal pair $(s, \hat{s})$ is then determined by the two likelihood ratios:

$$\frac{\mathbb{P}\left((s, \hat{s}) = (s_L, \hat{s_H}) \mid \theta = H\right)}{\mathbb{P}\left((s, \hat{s}) = (s_L, \hat{s_H}) \mid \theta = H\right)} = \frac{s_H}{1 - s_H} \tag{8.4}$$

$$\frac{\mathbb{P}\left((s, \hat{s}) = (s_L, \hat{s_L}) \mid \theta = H\right)}{\mathbb{P}\left((s, \hat{s}) = (s_L, \hat{s_L}) \mid \theta = H\right)} = \frac{s_L}{1 - s_L} \times \frac{1 - \varepsilon \times \dfrac{s_H}{1 - s_H}}{1 - \varepsilon \times \dfrac{s_L}{1 - s_L}} \tag{8.5}$$

Note that 8.5 is continuously and strictly decreasing with ε, taking values between $\frac{s_L}{1-s_L}$ and 0 as ε varies between 0 and $\frac{s_H}{1-s_H}$. The signal pair $(s, \hat{s})$ is informationally equivalent to $\mathcal{E}'$ when:

$$\frac{s_L}{1 - s_L} \times \frac{1 - \varepsilon \times \dfrac{s_H}{1 - s_H}}{1 - \varepsilon \times \dfrac{s_L}{1 - s_L}} = \frac{s_L - \delta}{1 - (s_L - \delta)}$$

I choose ε to satisfy this equality. As before, ε then becomes a continuously increasing function of δ.

Step 2: $\Pi(\sigma'_1; \mathcal{E}') - \Pi(\sigma_1; \mathcal{E}) \begin{cases} \geq 0 & s_L \geq s_L^{as} \\ \leq 0 & s_L \leq s_L^{as} \end{cases}$

The evaluator who observes the signal pair $(s, \hat{s})$ obtains equivalent information to that from $\mathcal{E}'$. We now must identify the strategy $\tilde{\sigma} : \{(s_L, \hat{s_H}), (s_L, \hat{s_L}), s_H\} \to [0, 1]$ for this signal pair that replicates the outcome of the strategy σ_1' for experiment $\mathcal{E}'$. This strategy is defined as:

$$\tilde{\sigma}(s_L, \hat{s_H}) = \tilde{\sigma}(s_H) = 1 \qquad\qquad \tilde{\sigma}(s_L, \hat{s_L}) = 0$$

and replicates the likelihood ratios of an approval and rejection signal under $\mathcal{E}'$.

Now, fix the applicant's *score profile*; the collection of signal draws evaluators will see for him were he to visit them all: $\mathbf{s} = \{(s^i, \hat{s}^i)\}_{i=1}^n$. I call an applicant a *marginal reject* if:

i for every $i \in \{1, 2, ..., n\}$, $s^i = s_L$, and

ii for at least one $i \in \{1, 2, ..., n\}$, $\hat{s}^i = \hat{s}_H$.

These marginal rejects drive the wedge between evaluators' payoffs under $\mathcal{E}'$ and $\mathcal{E}$: while *no* evaluator approves them under $\mathcal{E}$, *at least one* evaluator does under $\mathcal{E}'$. So:

$$\Pi(\sigma_1'; \mathcal{E}') - \Pi(\sigma_1; \mathcal{E}) = \mathbb{P}\left(\text{marginal reject}\right) \times \underbrace{\left[\mathbb{P}\left(\theta = H \mid \text{marginal reject}\right) - c\right]}_{(2)}$$

The evaluator of a marginal reject has either observed $(s^i, \hat{s}^i) = (s_L, \hat{s}_L)$, or $(s^i, \hat{s}^i) = (s_L, \hat{s}_H)$. Denote the latter number of evaluators as $\#$. Since the applicant is a marginal reject, $\# \geq 1$. Then, (2) equals:

$$\underbrace{\sum_{i=1}^{n} \frac{\mathbb{P}\left(i\ \hat{s_H}\text{ signals } \mid s^1 = ... = s^n = s_L\right)}{\sum_{j=1}^{n} \mathbb{P}\left(j\ \hat{s_H}\text{ signals } \mid s^1 = ... = s^n = s_L\right)}}_{(3)} \times \mathbb{P}\left(\theta = H \mid \# = i\right) - c$$

where:

$$\begin{aligned}\mathbb{P}\left(i\ \hat{s_H}\text{ signals } \mid s^1 = ... = s^n = s_L\right) = & k \times \binom{n}{i} \times \left(\frac{s_H}{1 - s_H} \times \varepsilon\right)^i \times \left(1 - \frac{s_H}{1 - s_H} \times \varepsilon\right)^{n-i} \\ & + (1-k) \times \binom{n}{i} \times \left(\frac{s_L}{1 - s_L} \times \varepsilon\right)^i \times \left(1 - \frac{s_L}{1 - s_L} \times \varepsilon\right)^{n-i}\end{aligned}$$

and $k = \mathbb{P}\left(\theta = H \mid s^1 = ... = s^n = s_L\right)$. Thus, the limit of expression (3) as $\varepsilon \to 0$ (and therefore, $\delta \to 0$) for any $i > 1$ is:

$$\lim_{\varepsilon \to 0} \frac{\frac{1}{\varepsilon} \times \mathbb{P}\left(i\ \hat{s} = \hat{s_H}\text{ signals } \mid s^1 = ... = s^n = s_L\right)}{\frac{1}{\varepsilon} \times \sum_{j=1}^{n} \mathbb{P}\left(j\ \hat{s} = \hat{s_H}\text{ signals } \mid s^1 = ... = s^n = s_L\right)} = 0 \tag{8.6}$$

Therefore, we get:

$$\lim_{\varepsilon\to 0}\sum_{i=1}^{n}\frac{\mathbb{P}\left(i\ \hat{s_H}\text{ signals }\mid s^1=...=s^n=s_L\right)}{\sum_{j=1}^{n}\mathbb{P}\left(j\ \hat{s_H}\text{ signals }\mid s^1=...=s^n=s_L\right)}\times\mathbb{P}\left(\theta=H\mid\#=i\right)-c$$

$$=\mathbb{P}\left(\theta=H\mid\#=1\right)-c$$

$$\propto\frac{\rho}{1-\rho}\times\left(\frac{s_L}{1-s_L}\right)^{n-1}\times\frac{s_H}{1-s_H}-\frac{c}{1-c}$$

proving the claim. □

□

Proposition 3. Suppose the experiment $\mathcal{E}$ is binary with outcomes inducing the normalised beliefs $s_L\in[0,0.5]$ and $s_H\in[0.5,1]$. Stronger bad news decrease evaluators' payoffs when:

i s_L is below the cutoff $\check{s}_L:=\min\{s_L^{\text{mute}},s_L^{as}\}$ in the least selective equilibrium.

ii s_L is below a cutoff $\hat{s_L}\geq\check{s_L}$ in the most selective equilibrium.

Proof. i The least selective equilibrium:

By Lemma 7, the probability that evaluators' approve upon the low outcome in the least selective equilibrium is:

$$\check{\sigma}^*(s_L)=\begin{cases}1 & s_L\geq s_L^{\text{mute}}\\ 0 & s_L<s_L^{\text{mute}}\end{cases}$$

Thus, the sum of their payoffs equals (i) the expected payoff from always approving the applicant when $s_L\geq s_L^{\text{mute}}$, and (ii) $\max\{0,\Pi(\sigma_1;\mathcal{E})\}$ when $s_L<s_L^{\text{mute}}$ (established in the proof of Theorem 1):

$$\Pi(\check{\sigma}^*;\mathcal{E})=\begin{cases}\rho-c & s_L\geq s_L^{\text{mute}}\\ \max\{0,\Pi(\sigma_1;\mathcal{E})\} & s_L<s_L^{\text{mute}}\end{cases}$$

Since approving all applicants is always feasible, we have $\max\{0,\Pi(\sigma_1;\mathcal{E})\}\geq\rho-c$ when $s_L<s_L^{\text{mute}}$ by Lemma 5. Furthermore, the final Claim in Theorem 1's proof establishes that as s_L falls, the expression $\max\{0,\Pi(\sigma_1;\mathcal{E})\}$ weakly increases (decreases) when $s_L\geq s_L^{\text{as}}$ ($s_L\leq s_L^{\text{as}}$). Thus the desired conclusion is established.

ii The most selective equilibrium:

By Lemma 8, the most selective equilibrium shifts from one where every applicant is approved to one where evaluators reject upon the low signal when $s_H\geq s_H^{\dagger}(s_L)$, where $s_H^{\dagger}(.)$ is an increasing function of s_L. Following the arguments made for the least selective equilibrium then, evaluators' most selective equilibrium payoffs:

- weakly increase as s_L decreases, when $s_L \geq \min\left\{s_L^{\text{as}}, s_L^{\dagger}(s_H)\right\}$
- weakly decrease as s_L decreases, when $s_L \leq \min\left\{s_L^{\text{as}}, s_L^{\dagger}(s_H)\right\}$.

The desired result follows by noting that $s_L^{\dagger}(s_H) \geq s_L^{\text{safe}}$, and therefore $\min\left\{s_L^{\dagger}, s_L^{\text{as}}(s_H)\right\} \geq \min\left\{s_L^{\text{safe}}, s_L^{\text{as}}(s_H)\right\}$.

□

Theorem 2. Say the experiment $\mathcal{E}'$ differs from $\mathcal{E}$ by a local mean preserving spread at s_j. Evaluators' payoffs under the most (least) selective equilibrium of $\mathcal{E}'$:

1. *weakly exceed* those under $\mathcal{E}$ if s_j leads to approvals under $\mathcal{E}$; i.e. $\hat{\sigma}_{\mathcal{E}}(s_j) = 1$ $(\check{\sigma}_{\mathcal{E}}(s_j) = 1)$
2. *fall weakly below* those under $\mathcal{E}$ if:

   i if s_j leads to rejections under $\mathcal{E}$; i.e. $\hat{\sigma}_{\mathcal{E}}(s_j) = 1$ $(\check{\sigma}_{\mathcal{E}}(s_j) = 1)$, and

   ii the following condition holds:

$$\frac{\rho}{1-\rho} \times \left(\frac{r_H(\sigma;\mathcal{E})}{r_L(\sigma;\mathcal{E})}\right)^{n-1} \times \frac{s_{j+1}}{1-s_{j+1}} \leq \frac{c}{1-c}$$

correct the r_H notation in thm statement

the notation shift from $\sigma_{\mathcal{E}}$ to σ and σ' confusing.

Proof. The Theorem focuses either on the least, or the most selective equilibrium strategies under both experiments. In the discussion below, I let σ^* and $\sigma^{*'}$ denote whichever equilibria we are focusing on under the respective experiments $\mathcal{E}$ and $\mathcal{E}'$. When I need to distinguish between the least and most selective equilibria, I denote them as $(\check{\sigma}, \check{\sigma}')$ and $(\hat{\sigma}, \hat{\sigma}')$, respectively. Following the notation introduced in Definition 4, let $\mathbf{S} \cup \mathbf{S}' = \{s_1, s_2, ..., s_M\}$ be the joint support of the experiments $\mathcal{E}$ and $\mathcal{E}'$, with elements increasing in their indices as usual. Since $\mathcal{E}'$ is obtained by a *local* mean preserving spread of $\mathcal{E}$, there is a monotone strategy $\sigma' : \mathbf{S}' \to [0,1]$ whose outcome under $\mathcal{E}'$ replicates the outcome of σ^* under $\mathcal{E}$:

$$\sigma'(s) = \begin{cases} \sigma^*(s_j) & s \in \{s_{j-1}, s_{j+1}\} \\ \sigma^*(s) & s \notin \{s_{j-1}, s_{j+1}\} \end{cases}$$

Claim 1. Evaluators' payoffs under the most (least) selective equilibrium of $\mathcal{E}'$ weakly exceed those under $\mathcal{E}$ when $\hat{\sigma}(s_j) = 1$ $(\check{\sigma}(s_j) = 1)$.

Now suppose s_j leads to approvals under σ^*; $\sigma^*(s_j) = 1$. Therefore, $\sigma'(s_{j-1}) = \sigma'(s_{j+1}) = 1$. Below, I show that $\sigma^{*'}$ is *more selective than* σ'. By Lemma 2, it follows that $\Pi\left(\sigma^{*'};\mathcal{E}'\right) \geq \Pi\left(\sigma';\mathcal{E}'\right) = \Pi\left(\sigma;\mathcal{E}\right)$.

If $s_{j-1} = \min \mathbf{S} \cup \mathbf{S}'$ or $\sigma^{*'}(s_{j-2}) = 0$, $\sigma^{*'}$ must necessarily be more selective than σ'; and we are done. So, for contradiction, I assume the following:

- $s_{j-1} > \min \mathbf{S} \cup \mathbf{S}'$

- ${\sigma^*}'(s_{j-2}) > 0$
- ${\sigma^*}'$ is *less* selective than σ', where the two strategies differ.

Case i. σ^* and ${\sigma^*}'$ are the least selective equilibrium strategies; i.e. $\sigma^* = \check{\sigma}$ and ${\sigma^*}' = \check{\sigma}'$.

I will prove the contradiction by constructing a strategy $\tilde{\sigma} : \mathbf{S} \to [0, 1]$ for experiment $\mathcal{E}$, such that:

i $\tilde{\sigma}$ replicates the outcome $\check{\sigma}'$ induces in $\mathcal{E}'$,

ii That $\check{\sigma}'$ is an eqm. strategy under $\mathcal{E}'$ implies that $\tilde{\sigma}$ is an eqm. strategy under $\mathcal{E}$,

iii But $\tilde{\sigma}$ is less selective than $\check{\sigma}$, contradicting that $\check{\sigma}$ is the least selective equilibrium strategy under $\mathcal{E}$.

I define the strategy $\tilde{\sigma} : \mathbf{S} \to [0, 1]$ for $\mathcal{E}$ as:

$$\tilde{\sigma}(s) := \begin{cases} 1 & s = s_i \\ \sigma'(s) & s \neq s_i \end{cases}$$

it is seen easily that $\tilde{\sigma}$ replicates the outcome of $\check{\sigma}'$. Furthermore, $\check{\sigma}'$ is an equilibrium under $\mathcal{E}'$ if and only if $\tilde{\sigma}$ is an equilibrium under $\mathcal{E}$: they induce the same interim belief ψ, and share the following necessary and sufficient condition for optimality:

$$\mathbb{P}_{\psi}\left(\theta = H \mid s_{j-2}\right) \begin{cases} = c & \sigma'(s_{j-2}) < 1 \\ \geq c & \sigma'(s_{j-2}) = 1 \end{cases}$$

The strategy $\tilde{\sigma}$ under experiment $\mathcal{E}$ replicates the outcome of $\check{\sigma}'$ under experiment $\mathcal{E}'$, and σ' under $\mathcal{E}'$ replicates the outcome of $\check{\sigma}$ under experiment $\mathcal{E}$. Since we assumed that $\check{\sigma}'$ is less selective than σ', it must be that $\tilde{\sigma}$ is less selective than $\check{\sigma}$.

i don't like the words "replicate the outcome". make more rigorous, add to definitions part.

Case ii. σ^* and ${\sigma^*}'$ are the most selective equilibrium strategies; i.e. $\sigma^* = \hat{\sigma}$ and ${\sigma^*}' = \hat{\sigma}'$.

Since strategy σ' for experiment $\mathcal{E}'$ replicates the outcome of $\hat{\sigma}$ for experiment $\mathcal{E}$, the two strategies induce the same interim belief ψ. Therefore, if $\mathbb{P}_{\psi}\left(\theta = H \mid s_{j-1}\right) \geq c$, σ' is an equilibrium under $\mathcal{E}'$; meaning $\hat{\sigma}'$ must be more selective than σ'.

Otherwise, say $\mathbb{P}_{\psi}\left(\theta = H \mid s_{j-1}\right) < c$. Then, by an argument used in the equilibrium locating algorithm I construct in insert this Lemma, there must be an equilibrium strategy that is more selective than σ' under $\mathcal{E}'$.

Claim 2. Evaluators' payoffs under the most (least) selective equilibrium of $\mathcal{E}'$ fall weakly below those under $\mathcal{E}$ if:

i. s_j leads to rejections under $\mathcal{E}$; i.e. $\hat{\sigma}(s_j) = 0$ ($\breve{\sigma}(s_j) = 0$), and

ii. the following condition holds:

$$\frac{\rho}{1-\rho} \times \left(\frac{r_H(\sigma;\mathcal{E})}{r_L(\sigma;\mathcal{E})}\right)^{n-1} \times \frac{s_{j+1}}{1-s_{j+1}} \leq \frac{c}{1-c}$$

Now, suppose s_j leads to rejections under σ^*; $\sigma^*(s_j) = 0$. Consequently, we have $\sigma'(s_{j-1}) = \sigma'(s_{j+1}) = 0$. I establish Claim 2 in two steps:

Step 1. $\sigma^{*'}$ is less selective than σ'; evaluators approve *more often* when s_j is mean preserving local spread.

Step 2. This decreases evaluators' payoffs when the condition in Claim 2 is met; $\Pi(\sigma^{*'};\mathcal{E}') \leq \Pi(\sigma';\mathcal{E}') = \Pi(\sigma^*;\mathcal{E})$.

Step 1.

If $s_{j+1} = \max \mathbf{S} \cup \mathbf{S}'$ or $\sigma^{*'}(s_{j+1}) > 0$, it must be the case that $\sigma^{*'}$ is less selective than σ', and we are done. So instead, I assume that $s_{j+1} < \max \mathbf{S} \cup \mathbf{S}'$ and $\sigma^{*'}(s_{j+1}) = 0$.

Case i. σ^* and $\sigma^{*'}$ are the least selective equilibrium strategies; i.e. $\sigma^* = \breve{\sigma}$ and $\sigma^{*'} = \breve{\sigma}'$

Since σ' replicates the outcome of $\breve{\sigma}$, we have $\Psi(\breve{\sigma};\mathcal{E}) = \Psi(\sigma';\mathcal{E}') = \psi$. Thus, σ' must be an equilibrium strategy under $\mathcal{E}'$ if $\mathbb{P}_\psi(\theta = H \mid s_{j+1}) \leq c$: the optimality conditions for all signals below s_{j+1} are satisfied *a fortiori*, and those for the signals above s_{j+1} are satisfied since $\breve{\sigma}$ has the same optimality conditions under $\mathcal{E}$. So, $\breve{\sigma}'$ must be less selective than σ', since the former is the least selective equilibrium.

If $\mathbb{P}_\psi(\theta = H \mid s_{j+1}, \psi) > c$ on the other hand, by an argument I used to locate equilibria in insert Lemma here, there must an equilibrium strategy that is *less* selective than σ' under experiment $\mathcal{E}'$.

Case ii. σ^* and $\sigma^{*'}$ are the most selective equilibrium strategies; i.e. $\sigma^* = \hat{\sigma}$ and $\sigma^{*'} = \hat{\sigma}'$.

$\hat{\sigma}'$ is the most selective equilibrium strategy under experiment $\mathcal{E}'$, and we assumed that $\hat{\sigma}'(s_{j+1}) = 0$. The strategy $\tilde{\sigma}$ defined below for experiment $\mathcal{E}$ replicates the outcome $\hat{\sigma}'$ generates under experiment $\mathcal{E}'$:

$$\tilde{\sigma}(s) = \begin{cases} 0 & s \leq s_j \\ \hat{\sigma}'(s) & s > s_j \end{cases}$$

Note that $\tilde{\sigma}$ must be an equilibrium under experiment $\mathcal{E}$, since the interim belief it induces is the same as the one $\hat{\sigma}'$ does, and its optimality constraints are a subset of the latter's. But since $\hat{\sigma}$ is the *most* selective equilibrium strategy under $\mathcal{E}$, $\tilde{\sigma}$ must be less selective than it.

Step 2.

The statement is trivially true when $\sigma' = \sigma^{*'}$, so I focus on the case where these two strategies differ. As Step 1 established, $\sigma^{*'}$ must be less selective than σ'. This implies that $\sigma^{*'}(s_{j+1}) > 0$. To see why, say we had $\sigma^{*'}(s_{j+1}) = 0$ instead. We can then construct a strategy $\tilde{\sigma}$ for experiment $\mathcal{E}$, which replicates the outcome $\sigma^{*'}$ generates under experiment $\mathcal{E}'$:

$$\tilde{\sigma}(s) = \begin{cases} 0 & s \leq s_j \\ \sigma^{*'}(s) & s > s_j \end{cases}$$

As they induce the same interim belief and the optimality constraints of the latter are a subset of the former's, $\tilde{\sigma}$ must be an equilibrium under $\mathcal{E}$. This contradicts with σ^* and $\sigma^{*'}$ being the least selective strategies; since $\sigma^{*'}$ being less selective than σ' implies that $\tilde{\sigma}$ must be less selective than σ^*. It also contradicts with σ^* and $\sigma^{*'}$ being the most selective strategies; since it would imply that σ', more selective than $\sigma^{*'}$, should be an equilibrium under $\mathcal{E}'$.

Given that $\sigma^{*'}(s_{j+1}) > 0$, I now take another strategy $\sigma^{\delta}_{\mathcal{E}'} : \mathbf{S}' \rightarrow [0, 1]$ for experiment $\mathcal{E}'$:

$$\sigma^{\delta}_{\mathcal{E}'}(s) = \begin{cases} 1 & s > s_{j+1} \\ \delta & s = s_{j+1} \\ 0 & s < s_{j+1} \end{cases}$$

where $\delta > 0$ is small enough so that $\sigma^{\delta}_{\mathcal{E}'}$ is more selective than $\sigma^{*'}$, but less selective than σ'. I will show that, when the condition stated in Claim 2 holds, we have $\Pi\left(\sigma^{\delta}_{\mathcal{E}'}; \mathcal{E}'\right) \leq \Pi\left(\sigma'; \mathcal{E}'\right)$ for $\delta \rightarrow 0$. Lemma 5 then implies that $\Pi\left(\sigma^{*'}; \mathcal{E}'\right) \leq \Pi\left(\sigma^{\delta}_{\mathcal{E}'}; \mathcal{E}'\right)$, which coins the result.

I show this slightly circuitously. I construct another experiment $\mathcal{E}^{\text{re}}$ under which I will use compare two strategies, σ_{re} and $\sigma^{\delta}_{\text{re}}$, that replicate the outcomes of the strategies σ' and $\sigma^{\delta}_{\mathcal{E}'}$, respectively. The experiment $\mathcal{E}^{\text{re}}$ has three possible outcomes, $\{s^{\text{re}}_L, s^{\text{re}}_{\delta}, s^{\text{re}}_H\}$. Conditional on the applicant's quality θ, its outcome distribution is independent from any other information any evaluator sees, and is given by the following pmf p^{re}_{θ}:

$$p_{\theta}(s^{\text{re}}) = \begin{cases} 1 - r_{\theta}(\sigma^*; \mathcal{E}) & s = s^{\text{re}}_H \\ \delta \times p'_{\theta}(s_{j+1}) & s = s^{\text{re}}_{\delta} \\ r_{\theta}\left(\sigma^*; \mathcal{E}\right) - \delta \times p'_{\theta}(s_{j+1}) & s = s^{\text{re}}_L \end{cases}$$

Define the strategies σ_{re} and $\sigma_{\text{re}}^{\delta}$ for this experiment as follows:

$$\sigma_{\text{re}}(s) = \begin{cases} 1 & s = s_H^{\text{re}} \\ 0 & s = s_\delta^{\text{re}} \\ 0 & s = s_L^{\text{re}} \end{cases} \qquad \sigma_{\text{re}}^{\delta} = \begin{cases} 1 & s = s_H^{\text{re}} \\ 1 & s = s_\delta^{\text{re}} \\ 0 & s = s_L^{\text{re}} \end{cases}$$

Now note that these two strategies replicate the outcomes of the strategies σ' and $\sigma_{\mathcal{E}'}^{\delta}$, respectively. Under $\sigma_{\text{re}}(s)$, the probability that an applicant is approved upon a visit, conditional on his quality, is the same as it is under strategy σ' (or σ^*, which it replicates), and under $\sigma_{\text{re}}^{\delta}$, it is the same as it is under $\sigma_{\mathcal{E}'}^{\delta}$.

So, the difference between evaluators' payoffs under these two strategies is determined by the *marginal reject* who:

- is rejected by *every* evaluator under the strategy σ_{re}.
- is approved by *at least one* evaluator under the strategy $\sigma_{\text{re}}^{\delta}$.

Where $\mathbf{s^{re}} = \left\{s^1, ..., s^n\right\}$ is the applicant's signal profile under the experiment $\mathcal{E}^{\text{re}}$, he has:

- *no* s_H^{re} signals; $s^i \neq s_H^{\text{re}}$ for all $i \in \{1, 2, ..., n\}$ and
- *at least one* s_δ^{re} signal; there exists some $i \in \{1, 2, ..., n\}$ such that $s^i = s_H^{\text{re}}$.

Thus we have:

$$\begin{aligned} \Pi(\sigma_{\mathcal{E}'}^{\delta}; \mathcal{E}') - \Pi(\sigma'; \mathcal{E}') &= \Pi(\sigma_{\text{re}}^{\delta}; \mathcal{E}^{\text{re}}) - \Pi(\sigma_{\text{re}}; \mathcal{E}^{\text{re}}) \\ &= \mathbb{P}\left(\text{marginal reject}\right) \times \underbrace{\left[\mathbb{P}\left(\theta = H \mid \text{marginal reject}\right) - c\right]}_{(2)} \end{aligned}$$

The expression labelled (2) above equals:

$$\sum_{i=1}^{n} \frac{\mathbb{P}\left(i\ s_\delta^{\text{re}} \text{ and } n-i\ s_L^{\text{re}} \text{ signals}\right)}{\sum\limits_{k=1}^{n} \mathbb{P}\left(k\ s_\delta^{\text{re}} \text{ and } n-k\ s_L^{\text{re}} \text{ signals}\right)} \times \mathbb{P}\left(\theta = H \mid i\ s_\delta^{\text{re}} \text{ and } n-i\ s_L^{\text{re}} \text{ signals}\right) - c$$

Since the probability that an evaluator observes the s_δ^{re} signal is proportional to δ, we have[13]:

$$\lim_{\delta \to 0} \frac{\mathbb{P}\left(i\ s_\delta^{\text{re}} \text{ and } n-i\ s_L^{\text{re}} \text{ signals}\right)}{\sum\limits_{k=1}^{n} \mathbb{P}\left(k\ s_\delta^{\text{re}} \text{ and } n-k\ s_L^{\text{re}} \text{ signals}\right)} = 0$$

[13]See expression 8.6 and the surrounding discussion in the proof of Theorem 1 for a more detailed explanation of this.

Therefore, we get:

$$\lim_{\delta\to 0}\sum_{i=1}^{n}\frac{\mathbb{P}\left(i\ s_\delta^{\mathrm{re}} \text{ and } n-i\ s_L^{\mathrm{re}} \text{ signals}\right)}{\sum_{k=1}^{n}\mathbb{P}\left(k\ s_\delta^{\mathrm{re}} \text{ and } n-k\ s_L^{\mathrm{re}} \text{ signals}\right)}\times\mathbb{P}\left(\theta=H \mid i\ s_\delta^{\mathrm{re}} \text{ and } n-i\ s_L^{\mathrm{re}} \text{ signals}\right)-c$$

$$\lim_{\delta\to 0}\mathbb{P}\left(\theta=H \mid \text{one } s_\delta^{\mathrm{re}} \text{ signal and } n-1\ s_L^{\mathrm{re}} \text{ signals}\right)-c$$

$$\propto\lim_{\delta\to 0}\frac{\rho}{1-\rho}\times\frac{p'_H(s_{j+1})}{p'_L(s_{j+1})}\times\left(\frac{r_H(\sigma^*;\mathcal{E})-\delta\times p'_H(s_{j+1})}{r_L(\sigma^*;\mathcal{E})-\delta\times p'_L(s_{j+1})}\right)^{n-1}-\frac{c}{1-c}$$

$$=\frac{\rho}{1-\rho}\times\frac{s_{j+1}}{1-s_{j+1}}\left(\frac{r_H(\sigma^*;\mathcal{E})}{r_L(\sigma^*;\mathcal{E})}\right)^{n-1}-\frac{c}{1-c}$$

□

Proposition 4. Let the experiment $\mathcal{E}'$ differ from $\mathcal{E}$ by a local mean preserving spread at s_j. Furthermore, let σ' and σ be the most selective equilibrium strategies under these two experiments. Evaluators' payoffs are lower with the strategies σ' if:

1. the signal s_j leads to rejections under σ; $\sigma(s_j)=0$, and
2. the following condition holds:

$$\frac{\rho}{1-\rho}\times\left(\frac{s_j}{1-s_j}\right)^{n-1}\times\frac{s_{j+1}}{1-s_{j+1}}\leq\frac{c}{1-c}$$

Proof. □

Lemma 3. Whenever it exists, the regulator-preferred garbling is *monotone binary* and *produces optimal recommendations.*

Proof. Add why optimal recommendations are without loss.

To prove this statement, I take an arbitrary garbling S^G and an equilibrium $\sigma^G:\mathbf{S}^G\to[0,1]$ it supports. I then construct a monotone binary garbling S^{G*} and equilibrium strategy σ^{G*} for it such that evaluators' payoffs are higher under σ^{G*} than under σ^G.

The garbling S^{G*} has signal realisations $\mathbf{S}^{G*}=\left\{s_L^{G*},s_H^{G*}\right\}$. It is monotone;

$$p_{s_k}^{G*}(s_L^{G*})>0 \quad\implies\quad p_{s_j}^{G*}(s_L^{G*})=1 \quad \text{for all } k>j$$

I define the strategy σ^{G*} for the garbling S^{G*} as:

$$\sigma^{G*}(s^{G*})=\begin{cases}0 & s^{G*}=s_L^{G*}\\ 1 & s^{G*}=s_H^{G*}\end{cases}$$

Under this strategy, the probability that a *Low* quality applicant is rejected from a visit is given by:

$$r_L^* := \sum_{j=1}^{m} \underbrace{p_L(s_j) \times p_{s_j}^{G*}\left(s_L^{G*}\right)}_{:=f^*(s_j)}$$

There is a unique set of probabilities $\{p_s^{G*}\}_{s\in\mathbf{S}}$ which (i) obeys the monotonicity constraint laid above, and (ii) where r_L^* equals r_L; the probability that an applicant with quality $\theta = L$ is rejected from a visit under the strategy σ^G for garbling S^G:

$$r_L^* = \sum_{j=1}^{m} \underbrace{p_L(s_j) \times \sum_{s^G\in\mathbf{S}^G} p_{s_j}^G(s^G) \times \left(1 - \sigma^G(s^G)\right)}_{:=f(s_j)}$$

I set $\{p_s^{G*}\}_{s\in\mathbf{S}}$ equal to these unique set of probabilities.

Now note that when normalised, we can treat the expressions f^* and f in the construction of r_L^* and r_L above as discrete probability densities over the set $\mathbf{S}$. Furthermore, the distribution the former describes is first order stochastically dominated by the one described by the latter; it's probability density $\frac{f^*(s_j)}{\sum_{s\in\mathbf{S}} f^*(s)}$ crosses $\frac{f(s_j)}{\sum_{s\in\mathbf{S}} f(s)}$ once from above. Therefore, we get:

$$\begin{aligned} r_H^* &:= \sum_{j=1}^{m} \frac{p_H(s_j)}{p_L(s_j)} \times \frac{f^*(s_j)}{\sum_{s\in\mathbf{S}} f^*(s)} \\ &\leq \sum_{j=1}^{m} \frac{p_H(s_j)}{p_L(s_j)} \times \frac{f(s_j)}{\sum_{s\in\mathbf{S}} f(s)} =: r_H \end{aligned}$$

where r_H^* and r_H are the probabilities that a *High* quality applicant is rejected from a visit under the strategies σ^{G*} and σ^G, respectively.

Since $r_H^* \geq r_H$ and $r_L^* = r_L$, evaluators' payoffs are higher under σ^* than they are under σ. It only remains to show that the strategy σ^* is optimal against the interim belief ψ^* consistent with it.

Thus, the interim belief ψ^* consistent with S^{G*} and σ^{G*} is below ψ, the one consistent with S^G and σ^G:

$$\frac{\psi^*}{1-\psi^*} = \frac{\sum_{k=0}^{n-1} (r_H^*)^k}{\sum_{k=0}^{n-1} \left(r_L^*\right)^k} = \frac{\sum_{k=0}^{n-1} (r_H)^k}{\sum_{k=0}^{n-1} (r_L)^k} \leq \frac{\psi}{1-\psi}$$

Under the interim belief ψ^*, it is optimal for an evaluator to reject an applicant upon the signal s_L^{G*} if and only if:

$$\frac{\psi^*}{1-\psi^*} \times \frac{r_H^*}{r_L^*} \leq \frac{c}{1-c}$$

But this inequality holds since $\frac{r_H^*}{r_L^*} \leq \frac{r_H}{r_L}$, $\psi^* \leq \psi$, and σ^G is optimal against the interim belief ψ:

$$\frac{\psi}{1-\psi} \times \frac{r_H}{r_L} \leq \frac{c}{1-c}$$

By Lemma XXX in Section XXX, $\Pi(\sigma^{G*}; S^{G*}) \geq \Pi\left(\sigma; S^G\right) \geq 0$ too. This then implies that approving applicants upon the signal s_H^{G*} must yield positive expected payoff. Hence, the strategy σ^{G*} is optimal against ψ^*.

□

Proposition 5. Let S^{G*} be the least selective monotone binary garbling with regret-free approvals, and σ^{G*} be the strategy that obeys its recommendations. Evaluators' payoffs under the strategy σ^{G*} and garbling S^{G*} exceed those under any other strategy and garbling pair.

Proof. It is without loss to compare S^{G*} only to monotone binary garblings S^G. Likewise, we can without loss fix evaluators' strategy σ^G to be:

$$\sigma^G(s^G) = \begin{cases} 0 & s^G = s_L^G \\ 1 & s^G = s_H^G \end{cases}$$

against any garbling S^G; evaluators' strategies need not satisfy any optimality constraints. The garblings S^{G*} and S^G replicate the strategies σ^* and σ over the original experiment $\mathcal{E}$, defined by $\sigma^*(s) := p_s^{G*}\left(p_H^{G*}\right)$ and $\sigma(s) := p_s^G\left(p_H^G\right)$. Note that:

$$a(s_H^G) = \min_{s \in \mathbf{S}} \sigma(s) > 0 \qquad\qquad r_\theta\left(\sigma; \mathcal{E}\right) = \sum_{s \in \mathbf{S}} p_L(s) \times p_s^G(s_L^G)$$

where the analogous equivalences hold between S^{G*} and σ^*.

Case 1: S^G is less selective than S^{G*}.

This implies σ is less selective than σ^*. Consider a third strategy σ^ε for the experiment $\mathcal{E}$ that is less selective than σ and such that $||\sigma^\varepsilon - \sigma|| = \varepsilon$ for some small $\varepsilon > 0$. Showing that $\Pi\left(\sigma^\varepsilon; \mathcal{E}\right) \leq \Pi\left(\sigma^*; \mathcal{E}\right)$ establishes that $\Pi\left(\sigma; \mathcal{E}\right) \leq \Pi\left(\sigma^*; \mathcal{E}\right)$ through Lemma XXX.

I showed when proving WHAT? that:

$$\lim_{\varepsilon \to 0} \Pi\left(\sigma^\varepsilon; \mathcal{E}\right) - \Pi(\sigma; \mathcal{E}) \propto$$

add the theorem 1 argument here!

Case 2: S^G is more selective than S^{G*}.

This implies σ is more selective than σ^*.

This is clearly worse for the planner because approving upon the worst signal, somebody with average rejections, is profitable.

□

Proposition 6. The regulator-preferred garbling is S^{G*} if its recommendations are optimal. Otherwise, it is either:

- the least selective monotone binary garbling with regret-free approvals, or
- the most selective monotone binary garbling with regret-prone approvals

among those with optimal recommendations.

Proof. □

Corollary 7. When evaluators' approval cost c is weakly above their prior ρ, S^{G*} is the regulator-preferred garbling.

Proof. □

References

Blackwell, D. (1953). Equivalent comparisons of experiments. The Annals of Mathematical Statistics, 24(2), 265–272. Retrieved October 4, 2022, from http://www.jstor.org/stable/2236332

Blackwell, D., & Girshick, M. A. (1954). Theory of games and statistical decisions. John Wiley & Sons.

Broecker, T. (1990). Credit-worthiness tests and interbank competition. Econometrica, 58(2), 429–452. Retrieved May 31, 2023, from http://www.jstor.org/stable/2938210

Gentzkow, M., & Kamenica, E. (2016). A rothschild-stiglitz approach to bayesian persuasion. The American Economic Review, 106(5), 597–601. Retrieved May 27, 2024, from http://www.jstor.org/stable/43861089

Kreyszig, E. (1978). Introductory functional analysis with applications. John Wiley & Sons.

Muirhead, R. F. (1900). Inequalities relating to some algebraic means. Proceedings of the Edinburgh Mathemati 19, 36–45. https://doi.org/10.1017/S0013091500032594

Müller, A., & Stoyan, D. (2002). Comparison methods for stochastic models and risks. Wiley. https://books.google.co.uk/books?id=a8uPRWteCeUC

Munkres, J. (2000). Topology. Prentice Hall, Incorporated. https://books.google.co.uk/books?id=XjoZAQAAIAAJ

Rasmusen, E., & Petrakis, E. (1992). Defining the mean-preserving spread: 3-pt versus 4-pt. In J. Geweke (Ed.), Decision making under risk and uncertainty: New models and empirical findings (pp. 53–58). Springer Netherlands. https://doi.org/10.1007/978-94-011-2838-4_7

Rothschild, M., & Stiglitz, J. E. (1970). Increasing risk: I. a definition. Journal of Economic Theory, 2(3), 225–243. https://doi.org/https://doi.org/10.1016/0022-0531(70)90038-4

Shaked, M., & Shanthikumar, J. (2007). Stochastic orders. Springer New York.