1 The Model

An *applicant* (he) seeks an approval from any one of $n \in \mathbb{N}_{\geq 2}$ *evaluators* (she), each with a distinct label $i \in \{1, 2, ..., n\}$. He is born with either *High* or *Low quality*; $\theta \in \{H, L\}$. Though θ is unknown both to himself[1] and the evaluators, all players correctly believe that he is born with *High* quality with probability $\rho \in (0, 1)$. To receive the approval he seeks, the applicant sequentially visits (applies to) all evaluators until either one *approves* him, or they all *reject* him. In the former case, I say he is *eventually approved*, and in the latter, *eventually rejected.* The order of his visits is described by the permutation $\tau(.)$ of the set of labels $\{1, 2, ..., n\}$. The applicant visits the evaluator labeled $\tau(i)$ after his $i-1^{\text{st}}$ rejection.

No evaluator knows the applicant's order of visits $\tau(.)$, but they commonly believe that they are all equally likely to be anywhere in the order; i.e. $\mathbb{P}(\tau(i) = j) = \frac{1}{n}$ for all $i, j \leq n$[2]. Thus, when an evaluator receives the applicant, she understands that he was rejected from all his earlier visits, but she does not know how many of these occurred. Therefore, she revises her prior belief ρ about the applicant's quality to an *interim belief* ψ_i upon receiving a visit. This revision is based on what she *believes* about the number of these past rejections and what they imply for the applicant's quality. I explain how evaluators form this interim belief in greater detail in Section 2.

Before evaluator i decides whether to *approve* the applicant she received, she observes the realisation of a private and costless signal x_i. x_i takes values in a finite set $S = \{s_1, s_2, ..., s_m\} \subset [0, 1]$. Without loss, I denote higher elements of S with higher indices: wherever $i > j$; $s_i \geq s_j$. Where I deal with *binary signals* with two possible realisations, $S = \{s_1, s_2\}$, it is friendlier to denote the low signal s_1 as s_L and the high signal s_2 as s_H instead. I do so.

x_i's distribution over S depends on the applicant's quality θ, and is described by the discrete density function $p_\theta : S \to [0, 1]$. Conditional on θ, evaluators' signals are IID; the distribution of an evaluator's signal depends neither on her label, nor on her order in τ. Without loss of generality, I label signal realisations after the *normalised posterior beliefs* they induce:

$$s = \frac{p_H(s)}{p_L(s) + p_H(s)} \qquad \text{wherever } p_H(s) + p_L(s) > 0$$

Likewise, I define the *normalised posterior density* p induced by $\mathcal{P}$ as:

$$p(s) := \frac{p_L(s) + p_H(s)}{2} \qquad \text{wherever } p_H(s) + p_L(s) > 0$$

Both the *actual* posterior beliefs and distribution $\mathcal{P}$ induces over them depend on evaluators' interim beliefs, themselves endogenous in this model. Generically, these coincide with their

[1] This is without loss of generality in the baseline model. The applicant's knowledge of his quality has no relevance to the analysis.

[2] I study the implications of relaxing this assumption later in Section INSERT .

normalised counterparts only when the interim belief assigns equal probability to either quality. Nonetheless, referring to the normalised posterior beliefs and density will be helpful especially when comparing the relative informativeness of signal structures.

is "generically" a right word to use here?

After observing the realised signal x_i, evaluator i decides whether to *approve* or *reject* the applicant. To approve the applicant, she incurs a fixed cost $c \in (0, 1)$; but whenever $\theta = H$, she also receives a benefit I normalise to 1. Upon an approval, the game ends and all other evaluators receive a payoff of 0. If the evaluator instead rejects her applicant, she receives a payoff of 0. The applicant then moves on to visit the evaluators he has not yet, unless none remain.

Evaluator i's *strategy* $\sigma_i : S \to [0, 1]$ maps every possible signal realisation $s \in S$ she might observe to a corresponding probability that she will approve the applicant, $\sigma_i(s)$. This strategy σ_i is *optimal* if:

$$\sigma_i(s) = \begin{cases} 0 & \mathbb{P}(\theta = H \mid \psi_i, x_i = s) < c \\ \in [0, 1] & \mathbb{P}(\theta = H \mid \psi_i, x_i = s) = c \\ 1 & \mathbb{P}(\theta = H \mid \psi_i, x_i = s) > c \end{cases}$$

where ψ_i is evaluator i's endogenously formed *interim belief* that the applicant has *High* quality, *given he visited her*.

interim belief? interim belief? what other name?

I focus on the *symmetric Bayesian Nash Equilibria* (henceforth just *equilibria*) of this game. An equilibrium is a strategy – belief pair (σ^*, ψ^*) such that:

1. each evaluator believes that $\theta = H$ with probability ψ^* given (i) the applicant visited her, and (ii) the strategies $\sigma_j = \sigma^*$ for all other evaluators $j \in \{1, 2, ..., n\}$,

2. the strategy σ^* is optimal given this belief ψ^*.

I call such a strategy σ^* an *equilibrium strategy*, and belief ψ^* an *equilibrium interim belief* under $\mathcal{P}$.

2 Interim Beliefs and Equilibria

Upon receiving a visit, an evaluator must assess the probability that she faces a *High* quality applicant. The signal x_i she observes is central to this assessment. However, she gleans crucial information about the applicant's quality from his mere visit, too.

The applicant visits our evaluator only if he was rejected by all other evaluators he visited earlier. Each such rejection itself is bad news about the applicant's quality. Our evaluator does not know how many such rejections occurred in the past. She nonetheless holds a belief about them. In particular, when all her peers use the strategy σ, she assigns a probability $r_\theta(\sigma; \mathcal{P})$ to the applicant receiving a rejection from any of his visits, conditional on having quality θ:

$$r_\theta(\sigma;\mathcal{P}) = 1 - \sum_{i=1}^{m} p_\theta(s_i)\sigma(s_i)$$

She believes she is – ex-ante – equally likely to be anywhere in the applicant's visit order $\tau(.)$. So she assigns a probability $\mathcal{R}_\theta(\sigma;\mathcal{P})$ to being visited before any of her peers approve the applicant:

$$\mathcal{R}_\theta(\sigma;\mathcal{P}) = \frac{1}{n} \times \sum_{k=1}^{n} r_\theta(\sigma;\mathcal{P})^{k-1}$$

The evaluator's interim belief ψ that the applicant who visits her has *High* quality is then a function $\Psi(.)$ of peer evaluators' strategies σ and the signal structure $\mathcal{P}$:

$$\psi = \Psi(\sigma;\mathcal{P}) := \frac{\rho \times \mathcal{R}_H(\sigma,\mathcal{P})}{\rho \times \mathcal{R}_H(\sigma,\mathcal{P}) + (1-\rho) \times \mathcal{R}_L(\sigma,\mathcal{P})}$$

When all evaluators use the strategy σ, they each hold the same interim belief ψ. This belief is determined *endogenously* in equilibrium; it depends on what evaluators' equilibrium strategies will be. However, the equilibrium strategies themselves must be optimal against the interim belief they induce. Thus, neither the existence nor the properties of equilibria are automatic.

Our first Proposition sets the ground by establishing some basic facts about equilibria. I exclude the uninteresting case of an uninformative signal structure, i.e. $p_H = p_L$. This is for brevity and without loss of interest: an equilibrium certainly exists in this case (either to approve *any* applicant, or *none*). In the knife edge case where we also have $\rho = c$, *any* strategy is an equilibrium; in any equilibrium evaluators are left indifferent between approving and rejecting the applicant regardless of the realised signal.

Proposition 1 first assures us that an equilibrium always exists. Further, any equilibrium strategy must be *monotone*: if a signal realisation s_i *might* lead to an approval, any better signal realisation $s_j > s_i$ *always* leads to an approval.

The monotonicty of equilibrium strategies critically implies that equilibria will always exhibit *adverse selection*. With a monotone strategy, a *Low* quality applicant is always likelier to get rejected than a *High* quality applicant. Thus, past rejections point to *Low* quality in any equilibrium. As no evaluator can rule out having possibly received the applicant after he has had many rejections already, evaluators view any visit to be *adversely selected*: an evaluator's interim belief ψ that the applicant who visited her has *High* quality is always weakly below her prior belief ρ.

Evaluators' interim beliefs are endogenous to their equilibrium strategies. So there might be multiple strategies which are optimal against the interim belief they induce; thus multiple equilibria. Nonetheless, Proposition 1 assures us that the set of equilibrium strategies is always *compact*. Compactness will be helpful to identify the *highest* and *lowest* equilibrium strategies, which play a special role in the sequel.

Equilibria will differ in the chances applicants stand to be approved. When $\sigma(s) \leq \sigma'(s)$ for every $s \in S$, or simply $\sigma \leq \sigma'$, I say σ is *more selective* than σ'. Conversely, I say σ' is *more embracive* than σ. A more selective (embracive) equilibrium offers applicants a higher probability of approval for any signal their evaluator might observe. Proposition 1 establishes that this ordering is *complete* over the set of equilibrium strategies: where σ and σ' both describe equilibrium strategy profiles, σ is either more selective or more embracive than σ'. This follows straightforwardly from the monotonicity of equilibrium strategies. In fact, this pointwise order is complete in the space of *all* monotone strategies.

should i retain "embracive" or not? "less selective" sounded derogatory.

Proposition 1. Let $p_H \neq p_L$. Where Σ is the set of equilibrium strategies:

1. *an equilibrium exists;* $\Sigma \neq \emptyset$,
2. *all equilibrium strategies are monotone;* for any $\sigma^* \in \Sigma$ and $s' > s$, $\sigma^*(s) > 0$ implies $\sigma^*(s') = 1$,
3. *all equilibria exhibit adverse selection;* $\psi^* \leq \rho$ for any ψ^* induced by an equilibrium strategy,
4. Σ is compact. Moreover, elements of Σ are pointwise totally ordered.

should i say anything about proof? quite standard.

Proof. See Section 4. □

Carlos: Easy to prove that more evaluators = more adverse selection *holding strategies fixed*. This out of eqm result didn't sound too interesting, so I excluded it for now.

That the set of equilibrium strategies is compact and totally ordered is important mainly because it implies that the *most* embracive and *most* selective equilibria, respectively the highest and lowest elements of this set, are well defined. I will focus heavily on these equilibrium strategies in the remainder of this paper, referring to them jointly as the *extreme equilibrium strategies*.

Notwithstanding its assurance that we can order equilibria from the most selective to the most embracive, Proposition 1 is silent about why this order would be useful. That gap is filled by Proposition 2. There, I establish that whether evaluators are better off in an equilibrium or the other is determined precisely by which one is more selective.

A priori, how moving towards more selective equilibria affects evaluators' payoffs is not clear. How evaluators fare in an equilibrium depends on how well they can distinguish and (i) eventually approve a *High* quality applicant, and (ii) eventually reject a *Low* quality one. In particular, when all evaluators use the strategy σ, the *sum* of all evaluators' expected payoffs are given by:

$$\begin{aligned}\Pi(\sigma;\mathcal{P}) := \rho \times (1-c) \times \mathbb{P}\left(\text{eventually approved} \mid \theta = H, \sigma_i = \sigma \ \forall i\right) \\ +(1-\rho) \times (-c) \times \left[1 - \mathbb{P}\left(\text{eventually rejected} \mid \theta = L, \sigma_i = \sigma \ \forall i\right)\right]\end{aligned} \tag{2.1}$$

$\Pi(\sigma;\mathcal{P})$ is just a sum of the probabilities that (i) *High* quality applicants are eventually approved, and (ii) *Low* quality applicants are eventually rejected. These probabilities are weighted by the the benefit of either outcome to evaluators, and the probability that the applicant was born with *High* quality. Note that we can easily recover a single evaluator's expected payoff $\pi_i(\sigma;\mathcal{P})$ from $\Pi(\sigma;\mathcal{P})$. $\pi_i(\sigma;\mathcal{P})$ is simply $\left(\frac{1}{n}\right)^{\text{th}}$ of $\Pi(\sigma;\mathcal{P})$ as all evaluators are ex-ante identical, and the equilibrium is symmetric.

Thus, selective and embracive equilibrium strategies have different virtues for evaluators' payoffs. Selective strategies depress applicants' chances of approval. By doing so, they protect evaluators against approving *Low* quality applicants too frequently. However, the high bar they impose for approvals potentially forsakes *High* quality applicants in the process. Vice versa, more embracive strategies give *High* quality applicants generous approval chances, but might be too admissive to *Low* quality applicants in the meanwhile.

Nonetheless, I establish in Corollary 1 that this trade-off is always resolved in favour of more selective equilibria. I show this by proving Proposition 2, which establishes an indeed stronger fact: evaluators' payoffs decrease *whenever* they move from an equilibrium strategy profile to *any* monotone strategy that's more embracive than it.

Proposition 2. Let σ^* be an equilibrium strategy, and σ be any other monotone strategy more embracive than σ^*. Evaluators' expected payoffs under σ^* exceed those under σ; $\Pi(\sigma^*;\mathcal{P}) \geq \Pi(\sigma;\mathcal{P})$.

Proof. See Section 4. $\square$

Corollary 1. Let σ^* and σ^{**} be two equilibrium strategies, where σ^{**} is more embracive than σ^*. Evaluators' expected payoffs under σ^* exceed those under σ^{**}; $\Pi(\sigma^*;\mathcal{P}) \geq \Pi(\sigma^{**};\mathcal{P})$.

The *most selective* and *most embracive* equilibria thus have special importance: they delineate the boundaries of payoffs evaluators can achieve across equilibria.

To grasp the intuition behind Proposition 2, consider moving all evaluators from an equilibrium strategy σ^* to a *marginally* more embracive strategy σ. Specifically, say σ^* and σ differ only for the highest signal which possibly leads to a rejection in σ^*, and that σ assigns only a marginally higher approval probability to this signal:

$$\sigma(\underline{s}) - \sigma^*(\underline{s}) > \varepsilon \qquad \text{where } \underline{s} := \min\{s \in S : \sigma^*(s) < 1\}$$

To further our thought experiment, fix the stream of signals $(x_1, x_2, ..., x_n)$ which evaluators will observe for this applicant if he visits them. Whether this applicant is eventually approved depends on what evaluators' strategies will instruct them to do upon the signal they observe.

Holding fixed this stream of signals, an applicant approved under the more selective strategy σ^* would also be approved under the more embracive strategy σ[3]. But the transition to σ might *also* lead to the eventual approval of some applicants who were rejected by every evaluator under σ^*. For such applicants, some of the signals their evaluators saw were only slightly below the mark under σ^* but still sufficient for an approval under σ.

is this foot-note clear? should i dwell on it?

With their eventual approval under σ, some such applicants might boost evaluators' expected payoffs. Consider for instance, an applicant who missed the mark only marginally in *every* evaluation he had. Under σ^*, all his evaluators lean on the cautious side, unable to rule out that he previously received multiple rejections, potentially with very low signals. If they knew that any previous rejections he had fell only marginally below the mark, they could have decided they were overestimating how adversely selected he was and revised their decision to an approval.

On the other hand, evaluators would rather keep some such applicants out, too. Consider, for instance, an applicant who faces a rejection by all but one of his evaluators under both strategy profiles. But under σ, say he scrapes through his *last* evaluation. Under σ^*, the bare *fear* of other evaluators' rejections, possibly with very low signals, had convinced this last evaluator to reject him. Her fear was in fact valid: the applicant was indeed rejected by all her peers. For all the worse, with signals that fell even below the more embracive standard σ imposes. Therefore, the approval of this applicant hurts evaluators' payoffs.

To prove Proposition 2, I identify which of these kinds of applicants the evaluators are likelier to approve on balance when their strategies get more embracive. I show that the probability of approving a *most* adversely selected applicant, of the latter kind, is overwhelmingly high. This is simply because when the difference between $\sigma(s_k)$ and $\sigma^*(s_k)$ shrinks, the probability that the applicant misses the mark under multiple evaluations vanishes rapidly. It thus becomes overwhelmingly likely that if he did miss the mark in some of his evaluations, he in fact did so in only *one*. He would fail all his other evaluations regardless of the strategy his evaluators used.

3 Equilibrium Payoffs with Better Informed Evaluators

As made evident by expression 2.1 for $\Pi(\sigma; \mathcal{P})$ in the previous Section, evaluators' payoffs are determined by how well they can distinguish between *High* and *Low* quality applicants. The extent of information their private signals x_i carry about θ play a critical role in this exercise.

[3]For simplicity, I sideline discussing possible differences in tie-breaking between σ^* and σ, which might of course mean an applicant approved under σ^* is rejected under σ; simply because he got unlucky when his evaluators were randomising in the latter. This is a distraction, and is easily remedied by appending evaluators' signals with a uniformly drawn tie-breaking signal according to which they break ties.

Indeed, if we had just one evaluator who was assessing the applicant for an approval, giving her a *(Blackwell) more informative* signal about θ would guarantee her a higher expected payoff[4]. It would do so by affording her a higher chance of approving a *High* quality applicant against *any* probability of rejecting a *Low* quality applicant that she wishes to target[5].

Armed with this classic insight, one might be hopeful that in our current setting too, evaluators would benefit from more information which would facilitate better evaluations of θ. Troublingly, this hope is unfounded. Zooming in to an individual evaluator i expected payoff $\pi_i(\sigma;\mathcal{P})$ when all evaluators use the strategies σ showcases what goes wrong:

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

The blue terms above, namely the probabilities that the evaluator (i) receives the applicant at all, and that (ii) his quality is *High* given the fact that she received him, are outside her control despite their direct relevance for her expected payoff. *When* the evaluator receives the applicant and *which* applicant she receives, thus the extent of *adverse selection* she faces, are shaped by *other* evaluators' equilibrium approval decisions. When they, now armed with more information, adjust their equilibrium decisions, the adverse selection inflicted on our evaluator might get aggravated. This might eclipse the better judgement more information affords her, leaving her worse off.

Here I am stretching the meaning of "adverse selection". Not just quality of applicant received but also when he is received.

This Section zooms in on this central question of this paper: how does better information about θ impact evaluators' expected payoffs, as it shapes and interacts with adverse selection?

3.1 With Binary Signals

Towards a general answer to this question, first consider evaluators who are restricted to *binary* signal structures; $S = \{s_L, s_H\}$. This case will build intuition and constitute the building block for the general answer I supply to this question in the next Section.

In a binary signal structure with signal realisations in $S = \{s_L, s_H\}$, the low signal s_L carries evidence of *Low* quality to the evaluator, and the high signal s_H of *High* quality. A binary signal structure $\mathcal{P}'$ is *(Blackwell) more informative* than another, $\mathcal{P}$, if its signals carry *stronger evidence* for both *High* and *Low* quality:

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

[4] Blackwell improving her signal is also is the *only* way to ensure this regardless of her prior belief ρ and approval cost c. The evaluator's new expected payoff is guaranteed to exceed what she could secure with her old signal. For completeness, I include a proof of this result in Lemma 2 in Section 4.2. This necessity owes to the evaluator facing a two state – two action decision problem.

[5] These, of course, owe to Blackwell, 1953's classic result. An exposition can be found in Blackwell and Girshick, 1954's Theorems 12.2.2 and Theorem 12.4.2.

where $\{s'_L, s'_H\}$ and $\{s_L, s_H\}$ are the possible signals an evaluator receives from $\mathcal{P}'$ and $\mathcal{P}$ respectively. Recall that these signals are labeled after the *normalised posterior beliefs* they induce. So one could express the condition above as:

$$\frac{p'_H(s'_L)}{p'_H(s'_L)+p'_L(s'_L)} \leq \frac{p_H(s_L)}{p_H(s_L)+p_L(s_L)} \qquad \frac{p'_H(s'_H)}{p'_H(s'_H)+p'_L(s'_H)} \geq \frac{p_H(s_H)}{p_H(s_H)+p_L(s_H)}$$

A more informative signal structure lets each evaluator discriminate better between *High* and *Low* quality applicants in her approval decisions. Figure INSERT illustrates this for binary signals:

Insert Figure and add explanation.

How do evaluators' equilibrium payoffs respond to stronger evidence for either *High* or *Low* quality; i.e. a signal with higher s_H or lower s_L? Theorem 1 establishes the answer, revealing that this depends precisely on *which* evidence gets stronger; i.e. *how* evaluators' information is improved. While stronger evidence for *High* quality always increases evaluators' equilibrium payoffs, stronger evidence for *Low* quality *eventually harms* them. Once it is too strong, providing evaluators even stronger evidence for *Low* quality can only hurt their equilibrium payoffs.

I removed the example in prev. draft.

Theorem 1. Let $\mathcal{P}$ be binary with normalised posterior beliefs $0 \leq s_L \leq 0.5$ and $0.5 \leq s_H \leq 1$. Evaluators' payoffs across the extreme equilibria are:

a) non-decreasing with the strength of evidence for $\theta = H$,

b) weakly increasing with the strength of evidence for $\theta = L$ when s_L is above a threshold,

c) weakly decreasing with the strength of evidence for $\theta = L$ when s_L is below that threshold,

for any approval cost $c \in [0, 1]$ and prior belief $\rho \in [0, 1]$.

Figure 1 illustrates Theorem 1 for the particular case where c is higher than the prior belief ρ. In that case, the equilibrium is unique for any signal structure. In that unique equilibrium, evaluators reject whenever they observe the low signal s_L. The figure plots the evolution of evaluators' payoffs across these unique equilibria. In the left panel, evidence for $\theta = L$ is strengthened as s_L falls from 0.5 to 0, for a fixed strength of evidence for $\theta = H$, s_H. Similarly in the right panel, evidence for $\theta = H$ is strengthened as s_H rises from 0.5 to 1 for fixed s_L.

Do I need a proof for this?

Figure 1 brushes equilibrium multiplicity away. Nonetheless, a key takeaway from Theorem 1 is that qualitatively, evaluators' equilibrium payoffs respond the same way to more information both across the most selective and the most embracive equilibria.

In the ensuing discussion, I sketch the main forces behind Theorem 1. I fix evaluators' strategies across signal structures to expose these forces: I let them simply approve upon s_H, and reject upon s_L. The full proof builds on this discussion, but deals carefully with equilibrium responses to changes in information, as well as possible equilibrium multiplicity. I relegate it to

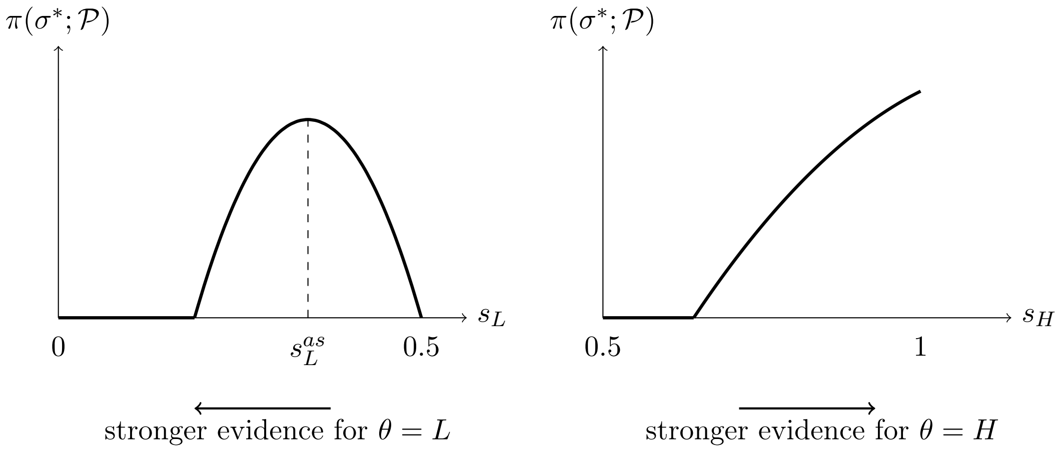


Figure 1: Equilibrium Payoffs with More Information, Binary Signals

Section 4. Theorem 1 does not comment on *where* the threshold after which stronger evidence for $\theta = L$ is harmful is. I characterise and discuss this threshold in Proposition 3, after sketching the main economic forces driving Theorem 1.

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."

Now consider replacing the signal structure $\mathcal{P}$ with a more informative one, $\mathcal{P}'$, which provides *marginally* stronger evidence for $\theta = L$, but not for $\theta = H$. In other words, $s'_L = s_L - \delta$ for some small $\delta > 0$ and $s'_H = s_H$, where $\{s'_L, s'_H\}$ are the signal realisations for $\mathcal{P}'$. Say evaluators approve an applicant whenever and only when they see a high signal realisation, under both signal structures. How do their payoffs react to this improvement in information?

The best way to see the answer is to consider giving evaluators the information $\mathcal{P}'$ provides with *two* signals x and $\hat{x}$. x comes from their old signal structure, $\mathcal{P}$. I design $\hat{x}$ to *complement* the information $\mathcal{P}$ provides to the one $\mathcal{P}'$ does. The evaluator observes the realisation of $\hat{x}$ *only* if her first observation is $x = s_L$. In turn, the new signal $\hat{x}$ is also binary, with possible realisations $\{\hat{s_L}, \hat{s_H}\}$. Conditional on θ, it is distributed independently from x with distribution:

"with distribution" right language?

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

ε, like δ, is a small positive number. I explain how it relates to δ in the next paragraph.

If the evaluator observes the signal $x = s_L$ followed by $\hat{x} = s_H$, her belief about θ jumps to what it would be if she saw $x = s_H$ in the first place. Note this from the likelihood ratio for this signal pair:

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

If he observes $\hat{x}=\hat{s_L}$ however, she loses further confidence that $\theta=H$. Specifically, the joint observation $(x,\hat{x})=(s_L,\hat{s_L})$ has the likelihood ratio:

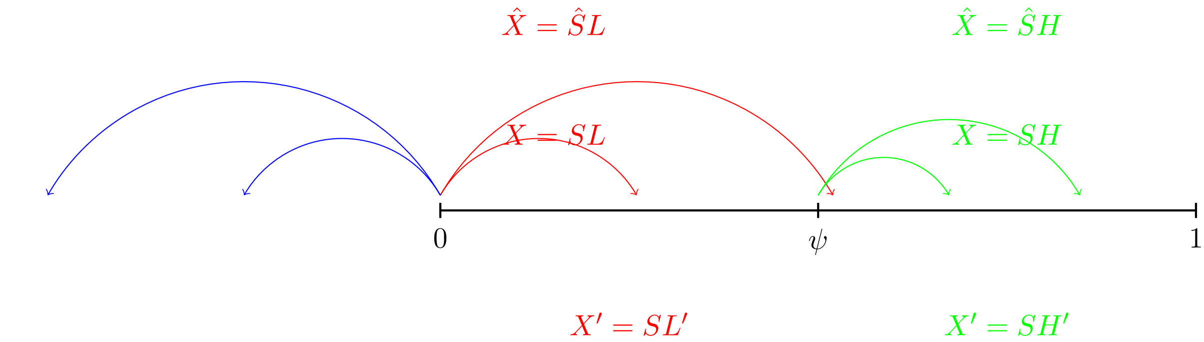
$$\frac{\mathbb{P}\left((x,\hat{x})=(s_L,\hat{s_L})\mid\theta=H\right)}{\mathbb{P}\left((x,\hat{x})=(s_L,\hat{s_L})\mid\theta=L\right)}=\underbrace{\frac{s_L}{1-s_L}\times\frac{1-\frac{s_H}{1-s_H}\times\varepsilon}{1-\frac{s_L}{1-s_L}\times\varepsilon}}_{(L,\hat{L})}<\frac{s_L}{1-s_L}$$

Note that the likelihood ratio $(L,\hat{L})$ above continuously and monotonically decreases as ε varies from 0 to $\frac{1-s_H}{s_H}$. We can thus choose ε to equate $(L,\hat{L})$ with the likelihood ratio for the signal $x'=s'_L$, or (L') below:

$$\frac{x'=s'_L\mid\theta=H}{x'=s'_L\mid\theta=L}=\underbrace{\frac{s_L-\delta}{1-(s_L-\delta)}}_{(L')}$$

Note that the value of ε which equates the likelihood ratios $(L,\hat{L})$ and (L') is then continuous and strictly increasing as a function of δ.

When the likelihood ratios $(L,\hat{L})$ and (L') are equal, the information an evaluator can extract about θ from the signal pair $(x,\hat{x})$ is equivalent to the information $\mathcal{P}'$ would provide her. Observing either $x=s_H$ or $(x,\hat{x})=(s_L,\hat{s_H})$ from the former carries the same information as observing $x'=s'_H$ from the latter. Similarly, the pair $(x,\hat{x})=(s_L,\hat{s_L})$ and the signal $x'=s'_L$ carry equivalent information. Thus for a given interim belief, the distribution of an evaluator's posterior beliefs about θ are the same whether she observes the signal tuple $(x,\hat{x})$ or just x'. This is illustrated in the left panel of Figure INSERT .



this figure will be edited

We wish our evaluator to approve all high signals and reject all low signals under $\mathcal{P}'$. Receiving the information contained in $\mathcal{P}'$ through the signal pair $(x,\hat{x})$ is no obstacle to implementing an equivalent decision rule. She simply approves *whenever* she sees a "positive" signal, $x=s_H$ or $\hat{x}=\hat{s_H}$, and rejects otherwise, upon $(x,\hat{x})=(s_L,\hat{s_L})$.

This reformulation of the new signal structure $\mathcal{P}'$ exposes *which* applicants' eventual outcomes change between $\mathcal{P}$ and $\mathcal{P}'$. Note that any applicant eventually approved under $\mathcal{P}$ would also be eventually approved under the pair $(x, \hat{x})$: the same $x = s_H$ signal suffices for an approval. However, an applicant who would be eventually rejected after n $x = s_L$ signals is given n "second chances" when the evaluators additionally observe $\hat{x}$. By overturning some of his initial $x = s_L$ signals with a $\hat{x} = \hat{s_H}$ signal, this applicant can gain back of some of his evaluators' favour, and overturn what *would* be a rejection to their approval.

How evaluators' expected payoffs change between $\mathcal{P}$ and $\mathcal{P}'$ depends on how likely this *marginal admit* has *High* quality. Any information available about his quality is contained in the collection of signal pairs $\{(x_i, \hat{x}_i)\}_{i=1}^n$ the evaluators would potentially observe for him, if he visited them all. Since he is a marginal admit, we know that initially all evaluators observed the signals $x = s_L$. How many of them later observed $\hat{x} = \hat{s_H}$?

For small δ, and therefore ε, the answer is *almost surely, one.* The probability that an evaluator observes $\hat{x} = \hat{s_H}$ is proportional to ε. So, that *multiple* evaluators observe $\hat{x} = \hat{s_H}$ has probability proportional to $o(\varepsilon)$. Therefore, as ε shrinks to 0, it becomes almost certain that the marginal admit overturned only *one* of her evaluators' negative assessments. Whether approving this marginal admit raises or decreases evaluators' payoffs then depends on whether one $\hat{x} = \hat{s_H}$ signal overwhelms $n-1$ $\hat{x} = \hat{s_L}$ signals in favour of the approval cost c:

$$\lim_{\delta \to 0} \frac{\rho}{1-\rho} \times \left(\frac{\hat{s_L}}{1-\hat{s_L}}\right)^{n-1} \times \frac{s_H}{1-s_H} = \frac{\rho}{1-\rho} \times \left(\frac{s_L}{1-s_L}\right)^{n-1} \times \frac{s_H}{1-s_H} \lesseqgtr \frac{c}{1-c}$$

The value of s_L at which the likelihood ratio above exactly equals $\frac{c}{1-c}$ will play a key role in our characterisation. I will call it s_L^{as}:

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

Beyond this tipping point, evaluators are worse off with any marginal admit *stronger* evidence for $\theta = L$ brings. This tipping point s_L^{as} has a special interpretation: it is the *strongest* evidence for $\theta = L$ can get before *adverse selection poses any threat whatsoever.* When evidence for $\theta = L$ is weaker than it, an evaluator observing $x = s_H$ would be happy to approve the applicant *even if* all her peers observed $x = s_L$ previously. Note that as well as the number of evaluators n, the approval cost c, and the unconditional prior ρ, s_L^{as} depends on how strong evidence for $\theta = H$ is, as well. The higher s_H, the lower s_L can get before adverse selection starts posing a threat.

Definition 1. For a given strength of evidence for *High* quality s_H, s_L^{as} is the *strongest* possible evidence for *Low* quality where adverse selection cannot pose a threat:

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

The *marginal admit* is precisely the applicant for whom *all bar one* evaluator observes a low signal. He determines how evaluators' payoffs change as we shift from $\mathcal{P}$ to $\mathcal{P}'$. The stronger evidence for $\theta = L$ gets, the more eager an evaluator would be to reject this marginal admit; if only she knew that the applicant was indeed him. Further lowering s_L can only strengthen this conviction, hurting evaluators' payoffs further.

In Figure 1, evaluators' equilibrium payoffs start decreasing precisely when s_L falls below s_L^{as}. This is no coincidence: in the case it illustrates, evaluators reject whenever they observe $x = s_L$, hence precisely the marginal admit I described above shapes their payoffs. The threshold after which stronger evidence for $\theta = L$ starts hurting evaluators' payoffs might in general differ from s_L^{as} due to equilibrium dynamics. Once I show why stronger evidence for $\theta = H$ benefits evaluators, I move to Proposition 3 where I explain and discuss this.

We can think of giving evaluators a signal structure $\mathcal{P}'$ with *marginally stronger* evidence for $\theta = H$, $s'_L = s_L$ but $s'_H = s_H + \delta$, analogously. This time evaluators observe realisation of the signal $\hat{x}$ *only if* they first receive $x = s_H$. The complementary signal $\hat{x}$ has the following distribution:

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

We choose ε to match the information $\mathcal{P}'$ provides with this signal pair:

$$\frac{\mathbb{P}\left((x,\hat{x}) = (s_H, \hat{s_H}) \mid \theta = H\right)}{\mathbb{P}\left((x,\hat{x}) = (s_H, \hat{s_H}) \mid \theta = H\right)} = \frac{s_H}{1-s_H} \times \frac{1-\varepsilon \times \frac{s_L}{1-s_L}}{1-\varepsilon \times \frac{s_H}{1-s_H}} = \frac{s_H+\delta}{1-s_H-\delta}$$

As before, ε is a continuous and increasing function of δ. I illustrate this construction in the right panel of Figure INSERT .

Upon observing the signal pair $(x, \hat{x})$, an evaluator's belief about θ either jumps up to where $x' = s'_H$ would take it, or down to where $x' = s'_L$ would. Thus, holding her interim belief constant, $(x, \hat{x})$ and $\mathcal{P}'$ induce the same distribution of posterior beliefs. An evaluator who would like to approve upon $x' = s'_H$ and reject upon $x' = s'_L$ can replicate this strategy with the pair $(x, \hat{x})$. She does so by approving only when she sees *two* high signals; $(x, \hat{x}) = (s_H, \hat{s_H})$. She rejects whenever her belief about θ sinks, either due to $x = s_L$, or to $\hat{x} = \hat{s_L}$.

This exposes the new group of applicants whose outcomes change as we transition from $\mathcal{P}$ to $\mathcal{P}'$. An applicant who would be rejected by all evaluators under $\mathcal{P}$ faces the same fate under $(x, \hat{x})$: the initial $x = s_L$ signals he received still lead to rejections. But an applicant who previously would be approved by an evaluator faces a renewed threat of rejection. His initial

$x = s_H$ signals can be overturned by $\hat{x} = \hat{s_H}$ signals in the second stage.

How evaluators' payoffs change between $\mathcal{P}$ and $\mathcal{P}'$ depends on how likely this *marginal reject* is to have *High* quality. As before, this information is contained in the collection of signal pairs his evaluators would observe for him, if he visited them all. Inferring these signals is now easier: as all eventually rejected him, *all* his evaluators must have seen low signals; either $x = s_L$ or $\hat{x} = \hat{s_L}$.

If a low signal *already* leads to a rejection in equilibrium, this marginal reject is *sure* to raise evaluators' expected payoffs. When evaluators find rejecting upon a low signal optimal, learning that *all* evaluators saw low signals can only strengthen this conviction.

Theorem 1 establishes that stronger evidence for $\theta = L$ hurts evaluators beyond a threshold, but does not uncover *where* this threshold is. Above I suggested s_L^{as}, the tipping point after which the *marginal admit* starts hurting evaluators' payoffs, as a natural candidate. However, whether the equilibrium effect of stronger evidence for $\theta = L$ is indeed described by this marginal admit depends very much on the equilibrium under scrutinty. When, for instance, *any signal* is being approved in this equilibrium anyway, strengthening evidence for $\theta = L$ can actually push the evaluators to a more selective equilibrium, thus raising their payoffs.

Nonetheless, equilibria where evaluators approve any applicant eventually disappear as their signal structure becomes more informative. In fact, once evidence for $\theta = L$ becomes strong enough to warrant a rejection *even if* the evaluator believed she was the *first one* to receive the applicant, no such equilibria can exist.

Definition 2. Let $\mathcal{P}$ be binary with normalised posterior beliefs $0 \leq s_L \leq 0.5$ and $0.5 \leq s_H \leq 1$. s_L^{mute} is the *lowest* value for s_L where approving *all* applicants is an equilibrium strategy:

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

Note that s_L^{mute} need not be below 0.5; in fact it never is when $c \geq \rho$.

Once such equilibria vanish, indeed evaluators' *marginal admits* determine how stronger evidence for $\theta = L$ affects their equilibrium payoffs. Proposition 3 lays out the threshold level of evidence for $\theta = L$ after which stronger such evidence harms evaluators, by taking these equilibrium dynamics into account.

Proposition 3. Let $\mathcal{P}$ be binary with normalised posterior beliefs $0 \leq s_L \leq 0.5$ and $0.5 \leq s_H \leq 1$. Evaluators' payoffs across the most embracive equilibria are:

i weakly increasing with the strength of evidence for $\theta = L$ when $s_L \geq \min\{s_L^{\text{mute}}, s_L^{as}\}$

ii weakly decreasing with the strength of evidence for $\theta = L$ when $s_L < \min\{s_L^{\text{mute}}, s_L^{as}\}$

Where $s_L^{\dagger} \geq \min\{s_L^{\text{mute}}, s_L^{as}\}$, their payoffs across the most selective equilibria are:

i weakly increasing with the strength of evidence for $\theta = L$ when $s_L \geq s_L^\dagger$

ii weakly decreasing with the strength of evidence for $\theta = L$ when $s_L < s_L^\dagger$

3.2 With General Discrete Signals

In the previous Section, we investigated how evaluators' equilibrium payoffs vary across binary signal structures. When her signal structure is binary, an evaluator observes either a high or a low signal; respectively evidencing *High* or *Low* quality. Such a signal structure is *more informative* whenever it provides *stronger evidence* in either direction. Theorem 1 showed that while stronger evidence for *High* quality always benefits evaluators, stronger evidence for *Low* quality *eventually* harms them.

In many settings of interest, however, the evaluators in concern have richer signal structures. Traders of financial assets and derivatives, for instance, might get recommendations at varying levels of strength, such as "Strong Sell", "Sell", "Buy", and "Strong Buy". This highlights the importance of extending our characterisation in Theorem 1 to improvements of such richer signal structures. In this Section, I characterise how receiving a more informative signal about quality affects evaluators' equilibrium payoffs for *any* discrete signal structure they might hold.

Evidence for this

I introduced a particular way to think about improving a binary signal structure in the previous Section. By carefully engineering an auxiliary signal which an evaluator sees only after a particular realisation of her original signal, we can replicate any improvement in her information. The auxiliary signal we engineer achieves this by spreading the evaluator's belief further, once she observes the initial evidence for *High* or *Low* quality. It is precisely by formalising and generalising this idea that I will characterise improvements to an *arbitrary* signal structure.

Definition 3 (Local Mean Preserving Spread)**.** Let p and p' be the normalised posterior densities for signal structures $\mathcal{P}$ and $\mathcal{P}'$. Additionally, let $\zeta_1 < \zeta_2 < ... < \zeta_M$ be the elements of $S \cup S'$; the joint support of $\mathcal{P}$ and $\mathcal{P}'$. If there exists some $i \in \{2, ..., M-1\}$ such that:

$$p'(\zeta_{i-1}) \geq p(\zeta_{i-1}) \qquad 0 = p'(\zeta_i) \leq p(\zeta_i) \qquad p'(\zeta_{i+1}) \geq p(\zeta_{i+1})$$

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

$$\sum_{i=1}^{M} \zeta_i \times p'(\zeta_i) = \sum_{i=1}^{M} \zeta_i \times p(\zeta_i)$$

I say $\mathcal{P}'$ *differs from* $\mathcal{P}$ *by a local mean preserving spread* (*at* ζ_i).

Much like an ordinary MPS (see Rothschild and Stiglitz, 1970)[6], a *local MPS* spreads proba-

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

The ζ notation is too confusing, just abuse notation

[6]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, MPS was first characterised by Muirhead, 1900 in the context of majorisation with *three* points. Rasmusen and Petrakis, 1992 show that these two ways of characterising MPS are in fact equivalent.

bility away from an *origin* point to two *destination* points: one above, and one below the origin point. It does so while preserving the mean of the original distribution. It differs crucially from an MPS however, in that the destination points must be the immediate neighbours of the origin point. In other words, the resulting distribution cannot assign positive probability to any point between the origin and the two destinations[7].

insert figure!

The auxiliary signal I introduced in the previous Section creates one such local MPS. To strengthen evidence for *Low* quality, for instance, the auxiliary signal we engineer spreads all the mass assigned to $x = s_L$ to the points $s_H > s_L$ and $s'_L < s_L$. Local mean preserving spreads are simple operations, but powerful enough to characterise *any* Blackwell improvement of a signal structure. $\mathcal{P}'$ is *Blackwell more informative than* $\mathcal{P}$, if and only if the former's normalised posterior distribution, F, can be transformed into the latter's, F', with finitely many local mean preserving spreads.

Remark 1. [Müller and Stoyan, 2002, Theorem 1.5.29] $\mathcal{P}'$ is Blackwell more informative than $\mathcal{P}$ if and only if there is a finite sequence $\mathcal{P}_1, \mathcal{P}_2, ..., \mathcal{P}_k$ such that $\mathcal{P}_1 = \mathcal{P}$, $\mathcal{P}_k = \mathcal{P}'$, and $\mathcal{P}_{i+1}$ differs from $\mathcal{P}_i$ by a local mean preserving spread.

We can thus decompose any improvement of a signal structure into finitely many local mean preserving spreads. It is these local mean preserving spreads whose effects I characterise in this Section.

this whole page needs thorough review and rewrite

As with binary signals, the effect of a local MPS will depend on whether it pushes evaluators to *approve* or *reject* more applicants. Spreads which create *rejects* always benefit evaluators. With the information such spreads contain, *any* initial favourable evaluation these applicants had is revised downwards. Eventually, thus, *all* evaluators judge it wiser to reject the applicant; even absent the knowledge of the unanimity of their verdict.

Spreads which create *admits*, however, can harm evaluators. Such spreads give applicants with no favourable evaluations the opportunity to overturn one of these and get approved by an evaluator. This evaluator, now with a positive assessment of the applicant's quality, approves him unaware that many of her peers are, or would be if they received the applicant, still pessimistic about his quality. This eventual approval ends up hurting evaluators' expected payoffs if the favourable information our evaluator received is overwhelmed by the remaining negative assessments of the applicant's quality.

DEFINITELY more explanation needed here, given how the proof goes.

Theorem 2 characterises how different spreads shape evaluators' payoffs by identifying which ones lead to rejects, and which ones to admits. This exercise is complicated by the fact that

[7]The attentive reader will also realise that this definition also requires a local MPS to spread *all* probability mass away from the origin point. This difference is insignificant for the results of this paper.

pinning down equilibria is, in general, a tedious task. Unlike with binary signals, we cannot straightforwardly categorise a signal realisation as "high" or "low"; how evaluators interpret it hinges on others' strategies, therefore on the signal structure it is nested in.

I adopt my approach to remedy this problem: locally spreading a signal which led to an approval in the original equilibrium must push evaluators towards more rejections. Similarly, locally spreading a signal which led to a rejection in the original equilibrium pushes evaluators towards more approvals.

here, too, more information on why local not arbitrary spreads are essential.

Before I finally state my main result, it will be helpful to describe the signals at which *adverse selection poses a threat*. For a given signal structure $\mathcal{P}$ and monotone strategy σ, I say *adverse selection poses a threat at signal* s, if an evaluator receiving this signal would like to reject the applicant *if she learned she was the last evaluator to be visited.*

Definition 4. Fix the signal structure $\mathcal{P}$ and a monotone strategy σ. I say *adverse selection poses a threat at signal* s if:

$$\frac{\rho}{1-\rho} \times \left(\frac{r_H(\sigma;\mathcal{P})}{r_L(\sigma;\mathcal{P})}\right)^{n-1} \times \frac{s}{1-s} \leq \frac{c}{1-c}$$

In words, if an evaluator receiving the signal s would like to *reject* the applicant *if she learned she was the last evaluator to be visited.*

As before, signals at which adverse selection poses a threat will be crucial to identify when a spread *does harm* evaluators.

Theorem 2. Let $\mathcal{P}'$ differ from $\mathcal{P}$ by a local MPS at s_i. Where σ' and σ both are either the most selective or most embracive equilibrium strategies under $\mathcal{P}'$ and $\mathcal{P}$, evaluators' expected payoffs under σ' are:

1. *weakly higher* than under σ, if $x = s_i$ leads to approvals under σ; $\sigma(\zeta_i) = 1$,

2. *weakly lower* than under σ, if:

   i $x = s_i$ leads to rejections under σ; $\sigma(\zeta_i) = 0$, and

   ii adverse selection poses a threat at signal ζ_{i+1}, for signal structure $\mathcal{P}$ and strategy σ.

obvious q: what about mixed? i think we can do something about it...

Proof. Following the notation introduced in Definition 3, let $S \cup S'$ be the joint support of the signal structures $\mathcal{P}$ and $\mathcal{P}'$, and $s_1 < s_2 < ... < s_M$ be its elements. I begin by noting that the outcome the monotone strategy $\sigma : S \to [0,1]$ generates under $\mathcal{P}$ can be replicated under $\mathcal{P}'$ by another monotone strategy $\tilde{\sigma}' : S' \to [0,1]$ provided $\sigma(\zeta_i) \in \{0,1\}$[8]:

$$\tilde{\sigma}'(s_j) = \begin{cases} \sigma(s_i) & j \in \{i-1, i+1\} \\ \sigma(s_j) & j \notin \{i-1, i+1\} \end{cases}$$

Part 1:

Now suppose s_i leads to approvals under σ; $\sigma(s_i) = 1$. Consequently, $\tilde{\sigma}'(s_{i-1}) = \tilde{\sigma}'(s_{i+1}) = 1$. I argue below that $\tilde{\sigma} \geq \sigma'$; evaluators *reject* more when s_i is spread. From Proposition 2, it follows that $\pi(\sigma; \mathcal{P}) = \pi(\tilde{\sigma}'; \mathcal{P}') \leq \pi(\sigma'; \mathcal{P}')$.

If $s_{i-1} = \min S \cup S'$ or $\sigma'(s_{i-2}) = 0$, we necessarily have $\tilde{\sigma} \geq \sigma'$ and are done. So, for contradiction, let $\sigma'(s_{i-2}) > 0$, and $\sigma' > \tilde{\sigma}'$.

Case 1: σ and σ' are the most embracive equilibria under $\mathcal{P}$ and $\mathcal{P}'$, respectively.

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

i $\tilde{\sigma}$ replicates the outcome σ' induces in $\mathcal{P}'$,

ii $\tilde{\sigma}$ is an equilibrium strategy under $\mathcal{P}$ if and only if σ' is an equilibrium strategy under $\mathcal{P}'$,

iii $\tilde{\sigma} > \sigma$, so σ cannot be the most embracive equilibrium under $\mathcal{P}$.

So, define the strategy $\tilde{\sigma} : S \to [0,1]$ for $\mathcal{P}$ as simply:

$$\tilde{\sigma}(s_j) := \begin{cases} 1 & j = i \\ \sigma'(s_j) & j \neq i \end{cases}$$

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

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

Lastly, since $\sigma' > \tilde{\sigma}'$, it must be that $\tilde{\sigma} > \sigma$.

Case 2: σ and σ' are the most selective equilibria under $\mathcal{P}$ and $\mathcal{P}'$, respectively.

Recall that $\tilde{\sigma}$ and σ induce the same interim belief ψ under their respective signal structures; $\Psi\left(\tilde{\sigma}; \mathcal{P}'\right) = \Psi\left(\sigma; \mathcal{P}\right) = \psi$. Therefore, if $\mathbb{P}\left(\theta = H \mid x = s_{i-1}, \psi\right) \geq c$, $\tilde{\sigma}$ is an equilibrium under $\mathcal{P}'$. This implies that $\sigma' \leq \tilde{\sigma}$, since σ' is the most selective equilibrium under $\mathcal{P}'$. If $\mathbb{P}\left(\theta = H \mid x = s_{i-1}, \psi\right) < c$ otherwise, there is an equilibrium $\sigma' < \tilde{\sigma}$ under $\mathcal{P}'$ due to the inter-

[8]The characterisation of $\tilde{}\sigma$ is otherwise the same, but it ceases to be *monotone* by our definition.

mediate value argument presented when equilibrium existence was established in Proposition 1.

make this reference clear, we are using it a lot.

Part 2:

Now suppose s_i leads to rejections under σ; $\sigma(s_i) = 0$. Consequently, $\tilde{\sigma}'(s_{i-1}) = \tilde{\sigma}'(s_{i+1}) = 0$. I will establish Theorem 2's claim in two steps:

1. $\sigma' \geq \tilde{\sigma}'$; evaluators *approve* more when s_i is spread,
2. $\pi(\sigma'; \mathcal{P}') \leq \pi(\tilde{\sigma}'; \mathcal{P}') = \pi(\sigma; \mathcal{P})$ when adverse selection poses a threat at signal s_{i+1} for signal structure $\mathcal{P}$ and strategy σ.

Step 1:

If $s_{i+1} = \max S \cup S'$ or $\sigma'(s_{i+1}) > 0$, we necessarily have $\sigma' \geq \tilde{\sigma}$. So, let $s_{i+1} < \max S \cup S'$ and $\sigma'(s_{i+1}) = 0$.

Case 1: σ and σ' are the most embracive equilibria under $\mathcal{P}$ and $\mathcal{P}'$, respectively.

Recall that $\Psi(\sigma; \mathcal{P}) = \Psi(\tilde{\sigma}'; \mathcal{P}') = \psi$ since $\tilde{\sigma}'$ replicates the outcome of σ. Thus, if:

$$\mathbb{P}(\theta = H \mid x = s_{i+1}, \psi) \leq c$$

$\tilde{\sigma}'$ must be an equilibrium strategy under $\mathcal{P}'$; the optimality conditions for all signals below s_{i+1} are satisfied *a fortiori*, and those for the signals above s_{i+1} are satisfied since σ is an equilibrium strategy in $\mathcal{P}$. Then, $\sigma' \geq \tilde{\sigma}'$, since σ' is the most embracive equilibrium. If

$$\mathbb{P}(\theta = H \mid x = s_{i+1}, \psi) > c$$

on the other hand, by the intermediate value argument we used to establish equilibrium existence in Proposition 1, there is an equilibrium strategy $\sigma' > \tilde{\sigma}$ under $\mathcal{P}'$.

Case 2: σ and σ' are the most selective equilibria under $\mathcal{P}$ and $\mathcal{P}'$, respectively.

Since $\sigma'(s_{i+1})$, its outcome under $\mathcal{P}'$ can be replicated $\mathcal{P}$ with a strategy $\tilde{\sigma}$, defined as:

$$\tilde{\sigma}(s_j) = \begin{cases} 0 & j = i \\ \sigma'(s_j) & j \neq i \end{cases}$$

I will show that this necessarily implies that $\sigma' \geq \tilde{\sigma}'$, in two steps:

i $\tilde{\sigma}$ is an equilibrium strategy under $\mathcal{P}$ if and only if σ' is an equilibrium strategy under $\mathcal{P}'$,

ii $\tilde{\sigma} \leq \sigma$, and therefore $\tilde{\sigma} = \sigma$ since σ is the most selective equilibrium under $\mathcal{P}$.

(i) follows trivially, since both strategies have the same optimality condition for every signal realisation above s_{i+1}. Now, since σ is the most selective equilibrium under $\mathcal{P}$, we must have

$\sigma \leq \tilde{\sigma}$; as $\tilde{\sigma}$ is an equilibrium strategy by (i). However, this means $\tilde{\sigma}' \leq \sigma'$. Since $\tilde{\sigma}'$ must also be an equilibrium in $\mathcal{P}'$, we must have $\tilde{\sigma}' = \sigma'$ and therefore $\sigma \leq \tilde{\sigma}$.

Step 2:

Now I establish the second part. The case where $\tilde{\sigma}' = \sigma'$ is trivial, so I focus on the case $\sigma' > \tilde{\sigma}'$. As we showed when establishing Case 2 in the first step, we must then have $\sigma'(s_{i+1}) > 0$.

Now take a strategy σ^{ε} for $\mathcal{P}'$, defined as $\sigma^{\varepsilon}(s_{i+1}) := \varepsilon$. We take ε small enough so that $\sigma' > \sigma^{\varepsilon} > \tilde{\sigma}'$. I will now show that when adverse selection poses a threat at signal s_{i+1} for $(\sigma; \mathcal{P})$, we have:

$$\pi(\sigma^{\varepsilon}; \mathcal{P}') \leq \pi(\tilde{\sigma}'; \mathcal{P}') = \pi(\sigma; \mathcal{P})$$

Proposition 2 then coins the result.

I show this slightly circuitously. Construct a ternary signal $\mathcal{P}^{\text{re}}$ which we will use to replicate the outcomes σ^{ε} and $\tilde{\sigma}'$ generate. This signal admits the realisations $x^{\text{re}} \in \{s_L^{\text{re}}, s_{\varepsilon}^{\text{re}}, s_H^{\text{re}}\}$ and has distribution:

$$\mathbb{P}\left(x^{\text{re}} = s \mid \theta\right) = \begin{cases} 1 - r_{\theta}(\sigma; \mathcal{P}) & s = s_H^{\text{re}} \\ \varepsilon \times p'_{\theta}(s_{i+1}) & s = s_{\varepsilon}^{\text{re}} \\ r_{\theta}(\sigma; \mathcal{P}) - \varepsilon \times p'_{\theta}(s_{i+1}) & s = s_L^{\text{re}} \end{cases}$$

Clearly, as defined below, the strategies σ^{re} and $\sigma^{\text{re-}\varepsilon}$ for $\mathcal{P}^{\text{re}}$ replicate the outcomes of $\tilde{\sigma}$ and σ^{ε} under $\mathcal{P}'$:

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

This makes it clear that the difference in evaluators' payoffs between $\tilde{\sigma}$ and σ^{ε} will be the *marginal admits* whose evaluators will observe:

i no s_H^{re} signal realisation,

ii at least one $s_{\varepsilon}^{\text{re}}$ signal realisation.

if they visit all evaluators. Thus, we have:

$$\Pi(\sigma^{\varepsilon}; \mathcal{P}') - \Pi(\tilde{\sigma}; \mathcal{P}') = \mathbb{P}\left(\text{marginal admits}\right) \times \underbrace{\left[\mathbb{P}\left(\theta = H \mid \text{marginal admit}\right) - c\right]}_{(2)}$$

where (2) then equals:

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

COMPLETE! $\square$

4 Proof Appendix

4.1 Useful Definitions and Notation

In what follows, I occasionally express beliefs in *likelihood form* for convenience. The reader can verify the following with ease: where $\frac{\psi}{1-\psi}$ is the likelihood ratio of an evaluator's belief about quality before he observes x_i, $\frac{s}{1-s} \times \frac{\psi}{1-\psi}$ is the likelihood ratio of her belief *after* observing $x_i = s$. It is (strictly) optimal for her to approve the applicant when this ratio (strictly) exceeds $\frac{c}{1-c}$.

Some strategies require evaluators to randomise when approving their applicant upon observing a particular signal realisation. To facilitate the discussion when so, I let evaluators formulate a *score* for the applicant they receive, based on the signal realisation x_i they observe.

Definition 5. An applicant's score from evaluator i is the pair (x_i, u_i), where $u_i \overset{IID}{\sim} U[0,1]$. His *score profile* $Z := \{(x_1, u_1), ..., (x_n, u_n)\}$ is the set of scores he would receive from each evaluator, were he to visit all.

Evaluator i approves the applicant upon his visit if and only if $\sigma_i(x_i) \leq u_i$.

Following the terminology of the applicant's *score profile*, I call the analogous set $\mathcal{X} := \{x_1, ..., x_n\}$ his *signal profile*.

4.2 Omitted Results

Lemma 2. Let there be a *single* evaluator who approves the applicant if and only if $\mathbb{P}(\theta = H \mid x) \geq c$. Her expected payoff under signal structure $\mathcal{P}'$ is greater than under $\mathcal{P}$ *regardless of her approval cost* $c \in (0,1)$ and prior belief $\rho \in (0,1)$ if and only if $\mathcal{P}' \underset{B}{\succeq} \mathcal{P}$.

Proof. Sufficiency is due to Blackwell's Theorem (Blackwell and Girshick, 1954, Theorem 12.2.2). I show necessity by fixing a prior belief ρ for the evaluator.

Let $F(.)$ and $F'(.)$ be the posterior distributions $\mathcal{P}$ and $\mathcal{P}'$ induce, respectively, for this prior belief ρ:

$$F(s_i) = (1-\rho) \times \sum_{j=1}^{i} p_L(s_j) + \rho \times \sum_{j=1}^{i} p_H(s_j)$$
$$F'(s_i) = (1-\rho) \times \sum_{j=1}^{i} p'_L(s_j) + \rho \times \sum_{j=1}^{i} p'_H(s_j)$$

The evaluator's expected payoff under $\mathcal{P}$ is then:

$$\int_c^1 (\omega - c)dF(\omega) = \int_c^1 \omega dF(\omega) - c \times (1 - F(c)) = (1-c) - \int_c^1 F(\omega)d\omega$$

Of course, an analogous expression gives her expected payoff under $\mathcal{P}'$. Therefore, for her expected payoffs under $\mathcal{P}'$ to exceed those under $\mathcal{P}$ for any $c \in (0,1)$, we must have:

$$\int_c^1 \left(F(\omega) - F'(\omega)\right) d\omega \geq 0$$

which is equivalent to $\mathcal{P}'$ being Blackwell more informative than $\mathcal{P}$[9]

□

4.3 Omitted Proofs

Proposition 1. Let $p_H \neq p_L$. Where Σ is the set of equilibrium strategies:

1. *an equilibrium exists;* $\Sigma \neq \emptyset$,
2. *all equilibrium strategies are monotone;* for any $\sigma^* \in \Sigma$ and $s' > s$, $\sigma^*(s) > 0$ implies $\sigma^*(s') = 1$,
3. *all equilibria exhibit adverse selection;* $\psi^* \leq \rho$ for any ψ^* induced by an equilibrium strategy,
4. Σ is compact. Moreover, elements of Σ are pointwise totally ordered.

Proof. Existence: I provide an algorithm to locate an equilibrium. Let $S = \{s_1, s_2, ..., s_m\}$, where signals with higher indices are higher without loss of generality. Again without loss, add two signals $s_0 = 0$ and $s_{m+1} = 1$ to the set (if $s_1 = 0$ or $s_m = 1$, replicate them). Where $e_i \in \mathbb{R}^{m+1}$ to be the standard basis vector who has 1 as its i^{th} entry, and 0 for all other entries, let $\sigma_j := \sum_{k=j}^{m} e_k$ be the strategy which approves the applicant if and only if the observed private signal exceeds the $j-1^{\text{st}}$ greatest signal in S. Denote the interim belief σ_j induces as $\psi_j := \Psi(\sigma_j, \mathcal{P})$. Note that the strategy σ_{m+1} never approves the applicant and $\psi_j = \rho$, as the signal realisation s_{m+1} is not in the support of $\mathcal{P}$.

If we have:

$$\frac{s_j}{1-s_j} \times \frac{\psi_j}{1-\psi_j} \geq \frac{c}{1-c} \geq \frac{s_{j-1}}{1-s_{j-1}} \times \frac{\psi_j}{1-\psi_j} \tag{4.1}$$

[9]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.

for some $j \in \{1, 2, ..., m+1\}$, then (σ_j, ψ_j) is an equilibrium. So say this holds for no index j. Then, we must have:

$$\frac{s_m}{1-s_m} \times \frac{\psi_{m+1}}{1-\psi_{m+1}} \geq \frac{s_m}{1-s_m} \times \frac{\psi_m}{1-\psi_m} > \frac{c}{1-c} > \frac{s_1}{1-s_1} \times \frac{\psi_1}{1-\psi_1}$$

The first inequality holds as $\psi_{m+1} = \rho \geq \psi_m$. The remaining part of the inequality follows from inequality 4.1 being violated both for $j = 1$ and $j = m+1$. Take the lowest index j^* such that:

$$\frac{s_{j^*}}{1-s_{j^*}} \times \frac{\psi_{j^*}}{1-\psi_{j^*}} \geq \frac{c}{1-c} > \frac{s_{j^*-1}}{1-s_{j^*-1}} \times \frac{\psi_{j^*-1}}{1-\psi_{j^*-1}}$$

Since Ψ is continuous in $\sigma(s_{j^*})$, we can find a strategy σ^* such that $\sigma_{j^*} \geq \sigma^* \geq \sigma_{j^*-1}$ which induces a interim belief ψ^* such that:

$$\frac{s_{j^*}}{1-s_{j^*}} \times \frac{\psi^*}{1-\psi^*} = \frac{c}{1-c}$$

and thus constitutes an equilibrium.

Monotonicity: Take some strategy σ that's optimal against the interim belief it induces $\psi = \Psi(\sigma; \mathcal{P})$. Optimality demands that $\sigma(s') \geq \sigma(s)$ for any $s', s \in S$ s.t. $s' > s$. Additionally, if $\sigma(s) \in (0, 1)$, it must be that:

$$\frac{s_H}{1-s_H} \times \frac{\psi}{1-\psi} > \frac{s}{1-s} \times \frac{\psi}{1-\psi} = \frac{c}{1-c} > \frac{s_L}{1-s_L} \times \frac{\psi}{1-\psi}$$

for any $s_H > s > s_L$, thus $\sigma(s_H) = 1$ and $\sigma(s_L) = 0$.

Mention once where this update comes from.

Adverse Selection: Take a *monotone* strategy σ. Since $p_H(.)$ likelihood ratio dominates $p_L(.)$, it also first order stochastically dominates it[10]. Therefore, $r_L(\sigma; \mathcal{P}) \geq r_H(\sigma; \mathcal{P})$ and $\Psi(\sigma; \mathcal{P}) \leq \rho$.

Totally Ordered and Compactness of the Equilibrium Set: Let Σ be the set of equilibrium strategies. Totally orderedness follows automatically since every element of Σ is a monotone strategy.

Since the set of *all* strategies is a bounded subset of $\mathbb{R}^m$, we only need to show that Σ is *closed* to establish compactness. So take a sequence $\{\sigma^n\} \in \Sigma$ s.t. $\sigma_n \to \sigma^*$. Note that all σ^n must be monotone strategies. Denote the respective interim beliefs as $\psi^n := \Psi(\sigma^n; \mathcal{P})$ and $\psi^* = \Psi(\sigma^*; \mathcal{P})$.

We would like to prove $\sigma \in \Sigma$. Wlog, I restrict attention to cases where $\{\sigma^n\}$ is either an increasing or a decreasing sequence; otherwise one can simply take a monotone subsequence of $\{\sigma^n\}$ which converges to the same limit. I take the case where $\{\sigma^n\}$ is a decreasing sequence here, the complementary case is analogous.

Let first $\sigma = \sigma_j$ for some $j \in \{1, 2, ..., m\}$. Then, there is some $N \in \mathbb{N}$ such that for all

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

$n \geq N$, $1 > \sigma^n(s_{j-1}) \to 0$ and thus:

$$\frac{\psi^n}{1-\psi^n} \times \frac{s_{j-1}}{1-s_{j-1}} = \frac{c}{1-c}$$

and so by the continuity of Ψ in $\sigma(s_{j-1})$:

$$\frac{\psi^*}{1-\psi^*} \times \frac{s_j}{1-s_j} > \frac{c}{1-c} = \frac{\psi^*}{1-\psi^*} \times \frac{s_{j-1}}{1-s_{j-1}}$$

thereby establishing that σ^* is an equilibrium strategy. The proof when $\sigma^*(s_j) \in (0,1)$ for some $j \in \{1, 2, ..., m\}$ is similar.

$\square$

The following Lemma will be of use when proving Proposition 2.

Lemma 3. Take three monotone strategies $\sigma'' > \sigma' > \sigma$. If $\Pi(\sigma'; \mathcal{P}) \leq \Pi(\sigma; \mathcal{P})$, then $\Pi(\sigma''; \mathcal{P}) \leq \Pi(\sigma'; \mathcal{P})$.

Proof. For the three strategies $\sigma'' > \sigma' > \sigma$, consider three sets $A, A', A'' \subset (S \times [0,1])^n$ where the applicant's score profile might lie:

$$\begin{aligned} &Z \in A && \text{if } Z \text{ is eventually approved under } \sigma'' \text{ but not } \sigma \\ &Z \in A' && \text{if } Z \text{ is eventually approved under } \sigma' \text{ but not } \sigma \\ &Z \in A'' && \text{if } Z \text{ is eventually approved under } \sigma'' \text{ but not } \sigma' \end{aligned}$$

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

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

and:

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

$\Pi(\sigma'; \mathcal{P}) \leq \Pi(\sigma; \mathcal{P})$ implies $\mathbb{P}(\theta = H \mid \mathbf{Z} \in A') \leq c$. But then we must have $\mathbb{P}(\theta = H \mid \mathbf{Z} \in A'') \leq c$, since $\mathbb{P}(\theta = H \mid Z \in A)$ is a convex combination of $\mathbb{P}(\theta = H \mid Z \in A')$ and $\mathbb{P}(\theta = H \mid Z \in A'')$, and:

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

Therefore, we have $\mathbb{P}(\theta = H \mid Z \in A'') \leq \mathbb{P}(\theta = H \mid Z \in A) \leq \mathbb{P}(\theta = H \mid Z \in A') \leq c$.

$\square$

Proposition 2. Let σ^* be an equilibrium strategy, and σ be any other monotone strategy more embracive than σ^*. Evaluators' expected payoffs under σ^* exceed those under σ; $\Pi(\sigma^*;\mathcal{P}) \geq \Pi(\sigma;\mathcal{P})$.

Proof. Let Z be the applicant's *score profile*, as in Definition 5. Take an equilibrium strategy σ^* and a more embracive strategy σ such that:

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

for some $\varepsilon > 0$, where $\underline{s} := \min\{s \in S : \sigma^*(s) < 1\}$. Now let $A \subset (S \times [0,1])^n$ be the set of score profiles which lead to rejections by all evaluators under σ^*, but an eventual approval under σ. Thus, $Z \in A$ if $x_i = \underline{s}$ and $\sigma(\underline{s}) \geq u_i > \sigma^*(\underline{s})$ for at least one evaluator $i \in \{1, ..., n\}$. Let $\#$ be the number of evaluators holding such scores for the applicant.

Only applicants' whose score profiles are in A change their eventual outcome from a rejection in σ^* to an approval in σ, thus:

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

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

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

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

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

I conclude the proof by showing that $\mathbb{P}(\theta = H \mid \# = 1) < c$, and invoking Lemma 3.

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

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

$$\frac{\psi^*}{1-\psi^*} = \frac{1 + r^*_H + ... + (r^*_H)^{n-1}}{1 + r^*_L + ... + (r^*_L)^{n-1}} \leq \left(\frac{r^*_H}{r^*_L}\right)^{n-1}$$

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

□

Theorem 1. Let $\mathcal{P}$ be binary with normalised posterior beliefs $0 \leq s_L \leq 0.5$ and $0.5 \leq s_H \leq 1$. Evaluators' payoffs across the extreme equilibria are:

a) non-decreasing with the strength of evidence for $\theta = H$,

b) weakly increasing with the strength of evidence for $\theta = L$ when s_L is above a threshold,

c) weakly decreasing with the strength of evidence for $\theta = L$ when s_L is below that threshold,

for any approval cost $c \in [0, 1]$ and prior belief $\rho \in [0, 1]$.

I will use the five lemmata below, possibly of independent interest, to prove Theorem 1. Throughout, I denote the most selective equilibrium under the signal structure $\mathcal{P}$ as $\sigma^{\text{sel}*}_{\mathcal{P}}$. Similarly, $\sigma^{\text{em}*}_{\mathcal{P}}$ is the most embracive equilibrium. The subscript is dropped when the signal structure in question is obvious.

insert numbers!

Lemma 4. Let $\mathcal{P}$ be binary. $\Psi(\sigma; \mathcal{P})$ 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 $\sigma(F) \in (0, 1)$ and $\sigma(S) = 1$. The interim belief ψ is then given by:

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

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:

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

The proof is completed by noting that:

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

decreases, and thus ψ increases, in $\sigma(s_L)$.

Part ii:

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

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

and:

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

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. Thus, $\Psi(\sigma;\mathcal{P})$ decreases in $\sigma(s_H)$.

$\square$

The Corollary below follows from Lemma 4: if $\mathcal{P}' \underset{B}{\succeq} \mathcal{P}$ where both signal structures are binary, adverse selection is stronger under $\mathcal{P}'$, if evaluators always (i) approve upon the high signal, and (ii) reject upon the low signal, under both signal structures.

Corollary 5. Let $\mathcal{P}'$ be more informative than $\mathcal{P}$, the strategies $\sigma'_{(0,1)}$ and σ_1 be $\sigma'_{(0,1)}(s'_L) = \sigma_1(s_L) = 0$ and $\sigma'_{(0,1)}(s'_H) = \sigma_1(s_H) = 1$. Then, $\Psi(\sigma';\mathcal{P}') \leq \Psi(\sigma;\mathcal{P})$.

Proof. I will only prove that the assertion holds when $s_L = s'_L$ but $s'_H > s_H$. The mirror case, which establishes the second part of the corollary, is analogous.

The proof will show that the outcome induced by σ under signal structure $\mathcal{P}$ can be replicated by $\tilde{\sigma}$ under signal structure $\mathcal{P}'$, where $\tilde{\mathcal{P}}(s_L) > 0$ and $\tilde{\mathcal{P}}(s_H) = 1$. Then, the conclusion follows from Lemma 4.

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

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

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

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

where the family of distributions $\{p'_\theta\}$ belong to $\mathcal{P}'$. It is easy to verify that this last fraction 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 $(\sigma; \mathcal{P})$.

This proves the corollary.

$\square$

Lemma 6. Let $\mathcal{P}$ be binary. There is no mixing at $x = s_L$ neither in $\sigma^{\text{sel}*}$ nor in $\sigma^{\text{emb}*}$; i.e. $\sigma^{\text{sel}*}(s_L), \sigma^{\text{em}*}(s_L) \in \{0, 1\}$.

Proof. I start by showing $\sigma^{\text{em}*}(s_L) \in \{0, 1\}$. Where s_L^{mute} is as it was defined in Definition 2, observe that when $s_L \geq s_L^{\text{mute}}$, $\sigma(s_L) = \sigma(s_H) = 1$ is an equilibrium. This is because $\psi = \rho$ at this induced equilibrium, thus approving upon the low signal is optimal. This is the most embracive equilibrium, since there is no strategy that's more embracive. When $s_L < s_L^{\text{mute}}$, any equilibrium σ must feature $\sigma(s_L) = 0$, since $\psi \leq \rho$.

Now consider $\sigma^{\text{sel}*}$. For contradiction, let $\sigma^{\text{sel}*}(s_L) > 0$. By Lemma 4, and an argument used while proving equilibrium existence in Proposition 1, there is then another equilibrium σ where $\sigma(s_L) = 0$. clarify!

$\square$

Lemma 7 characterises evaluators' decisions upon seeing the low signal in the extreme equilibria, given how informative $\mathcal{P}$ is. Broadly, more informative signal structures push evaluators to reject upon the low signal under both equilibria.

Lemma 7. Let $\mathcal{P}$ be binary, with signal realisations s_L and s_H. Then:

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

where $s_H^{\dagger}(.)$ is an increasing function, and $s_H^{\dagger}(s_L^{\text{mute}}) = 0.5$.

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 proves the part for the most embracive equilibrium, combined with Lemma 6.

Let the strategy σ_1 be such that $\sigma_1(s_L) = 0$ and $\sigma_1(s_H) = 1$. The following is a necessary and sufficient condition for an equilibrium σ^* where $\sigma^*(s_L) = 0$ to exist is:

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

Necessity follows from $\Psi(\sigma_1; \mathcal{P}) \geq \Psi(\sigma^*; \mathcal{P})$ due to Lemma 4. Sufficiency follows from the fact that an equilibrium always exists, and the condition above implies s_L must always be rejected in it. Due to Corollary 5, we know that this condition holds when s_H is weakly above some threshold $s_H^{\dagger}(s_L)$, increasing with s_L. The necessary and sufficient condition holds whenever s_L^{mute}, therefore $s_H^{\dagger}(s_L^{\text{mute}}) = 0.5$.

□

Proof, Theorem 1: I prove Theorem 1 by establishing four facts:

1. The expected payoff in an equilibrium σ^* where $\sigma^*(s_L) = 0$ is higher than the expected payoff of approving all applicants.

   This follows directly from Proposition 2.

2. There is at most one equilibrium where $\sigma^*(s_L) = 0$.

   Let $\{\sigma_\alpha\}_{\alpha \in [0,1]}$ be the family of strategies where the low signal is rejected: $\sigma_\alpha(s_L) := 0$ and $\sigma_\alpha(s_H) := \alpha$. If:

   $$\frac{\Psi(\sigma_1; \mathcal{P})}{1-\Psi(\sigma_1; \mathcal{P})} \times \frac{s_H}{1-s_H} \geq \frac{c}{1-c}$$

   σ_1 is the only equilibrium candidate among this family; the interim belief is higher under any lower α by Lemma 4. Otherwise, again by Lemma 4, there is at most one $\alpha \in [0,1]$ for which:

   $$\frac{\Psi(\sigma_\alpha; \mathcal{P})}{1-\Psi(\sigma_\alpha; \mathcal{P})} \times \frac{s_H}{1-s_H} - \frac{c}{1-c} = 0$$

When such an α exists, σ_α is the only equilibrium candidate in this family. Under higher α, approving upon $x = s_H$ is not optimal. Under lower α, rejecting upon $x = s_L$ is not optimal. If the expression above is strictly negative for *any* α, then the only equilibrium candidate where the low signal is rejected is σ_0.

3. When an equilibrium $\sigma^* \in \{\sigma_\alpha\}_{\alpha \in [0,1]}$ where all low signals are rejected exists, the expected payoff in this equilibrium is given by $\pi_i(\sigma^*; \mathcal{P}) = \max\{0, \pi_i(\sigma_1; \mathcal{P})\}$.

   Above we showed that evaluators expect positive expected payoff (necessarily from approving an applicant) only when $\alpha = 1$. Otherwise, they either approve no applicant or are indifferent to rejecting those they do.

Theorem 1 then follows from our fourth claim:

Below, I use s_L^{as} but do not reiterate what it means.

make sure the notation is either Π or π

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

   i weakly increasing in s_H whenever σ_α is an equilibrium strategy for some $\alpha \in [0, 1]$,

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

   - weakly increasing when $s_L \geq s_L^{as}$,
   - weakly decreasing when $s_L \leq s_L^{as}$

   where s_L^{as} is implicitly defined as:

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

   for the signal structure $\mathcal{P}$.

Due to Lemma 7, both the most embracive and most selective equilibria shift once from the equilibrium where *all* s_L signals are approved to the one where *none* are approved, as the binary signal structure $\mathcal{P}$ becomes more informative. Due to the first fact laid out in the proof of this Theorem, this induces an increase in evaluators' expected payoff. Therefore, this last assertion about the shape of evaluators' payoffs in the equilibrium where the low signal is rejected concludes the proof.

Proof for the fourth claim:

Part i: Increasing s_H; i.e. the strength of evidence for $\theta = H$.

Let $\mathcal{P}$ and $\mathcal{P}'$ be two binary signal structures with respective signal realisations $\{s_L, s_H\}$ and $\{s'_L, s'_H\}$. Let $s'_L = s_L$ and $s'_H = s_H + \delta$ for $1 - s_H \geq \delta > 0$. I show that $\Pi(\sigma_1; \mathcal{P}') > \Pi(\sigma_1; \mathcal{P})$.

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

Rather than having evaluators observe one draw from the signal structure $\mathcal{P}'$, say an evaluator potentially observes *two* signal realisations; x and $\hat{x}$. She first observes x, a single draw from $\mathcal{P}$. If this signal realises as $x = s_L$, she observes no further information. If instead $x = s_H$, she observes another signal $\hat{x} \in \{\hat{s_L}, \hat{s_H}\}$, a draw from the signal structure $\hat{\mathcal{P}}$. $\hat{x}$ has the following distribution, and is independent from x, conditional on θ:

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

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

$$\frac{\mathbb{P}((x,\hat{x}) = (s_H, \hat{s_H}) \mid \theta = H)}{\mathbb{P}((x,\hat{x}) = (s_H, \hat{s_H}) \mid \theta = L)} = \frac{s_H}{1 - s_H} \times \frac{1 - \varepsilon \times \dfrac{s_L}{1 - s_L}}{1 - \varepsilon \times \dfrac{s_H}{1 - s_H}} \tag{4.2}$$

$$\frac{\mathbb{P}((x,\hat{x}) = (s_H, \hat{s_L}) \mid \theta = H)}{\mathbb{P}((x,\hat{x}) = (s_H, \hat{s_L}) \mid \theta = L)} = \frac{s_L}{1 - s_L} \tag{4.3}$$

Note that the likelihood ratio 4.2 increases continuously with ε. The signal pair $(x, \hat{x})$ is informationally equivalent to $\mathcal{P}'$ when:

$$\frac{s_H}{1 - s_H} \times \frac{1 - \varepsilon \times \dfrac{s_L}{1 - s_L}}{1 - \varepsilon \times \dfrac{s_H}{1 - s_H}} = \frac{s_H + \delta}{1 - (s_H + \delta)} \tag{4.4}$$

for our chosen (δ, ε). Choose ε to satisfy this equality; note that ε becomes a continuously increasing function of δ. Furthermore, by varying ε between 0 and $\frac{1-s_H}{s_H}$, the equivalent of *any* signal structure $\mathcal{P}'$ with $s'_L = s_L$ and $1 \geq s'_H \geq s_H$ can be obtained.

Step 2: $\pi(\sigma_1; \mathcal{P}') > \pi(\sigma_1; \mathcal{P}')$.

The strategy σ_1 can be replicated by an evaluator who receives the signal pair $(x, \hat{x})$ instead of x'. To do so, the evaluator approves if and only if the pair $(x, \hat{x}) = (s_H, \hat{s_H})$ is observed. Note that, conditional on the visiting applicant's quality, the probability that the evaluator approves him is the same under these two policies. This is due to the identical informational content of these signals, as laid out in equations 4.3 and 4.4. Thus, evaluators' payoffs are also identical under these policies.

Fix the collection of signal draws evaluators will see for the applicant if he visits them all: $\{(x_i, \hat{x}_i)\}_{i=1}^n$. An applicant is a *marginal reject* if he has no $(x_i, \hat{x}_i) = (s_H, \hat{s_H})$ signals. The difference between evaluators' payoffs under $(\mathcal{P}, \hat{\mathcal{P}})$ and $\mathcal{P}$ is determined by these *marginal rejects*: they are *eventually rejected* under $(\mathcal{P}, \hat{\mathcal{P}})$ but *eventually approved* under $\mathcal{P}$. So:

$$\Pi(\sigma_1; \mathcal{P}') - \Pi(\sigma_1; \mathcal{P}) = \mathbb{P}\,(\text{marginal reject}) \times \underbrace{[c - \mathbb{P}\,(\theta = H \mid \text{marginal reject})]}_{(1)}$$

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

$$c - \mathbb{P}\left[\theta = H \mid x_1 = ... = x_n = s_L\right]$$

In the relevant region where $x = s_L$ leads to a rejection, the expression above must be weakly positive. Therefore, $\Pi(\sigma_1; \mathcal{P}') - \Pi(\sigma_1; \mathcal{P}) \geq 0$.

Maybe a brief explainer.

This concludes the first part of the claim that $\max\{0, \pi(\sigma_1; \mathcal{P})\}$ is weakly increasing in s_H.

I should probably just focus on the π not the whole thing, correct!

Part ii: Decreasing s_L; i.e. increasing the strength of evidence for $\theta = L$.

Now I show that replacing $\mathcal{P}$ with $\mathcal{P}'$ when $s'_L = s_L - \delta$ and $s'_H = s_H$:

i increases $\pi(\sigma_1; \mathcal{P})$ when $s_L \leq s_L^{as}$,

ii decreases $\pi(\sigma_1; \mathcal{P})$ when $s_L > s_L^{as}$

for $\delta > 0$ arbitrarily small. The desired assertion follows.

I should probably use the sum of evaluators' payoffs here.

Step 1: Replicating $\mathcal{P}'$ in two signals.

As before, let the evaluator potentially observe *two* signal realisations, x and $\hat{x}$. She first observes x, a single draw from $\mathcal{P}$. If this signal realises as $x = s_H$, she receives no further information. If it realises as $x = s_L$, she observes another signal $\hat{x} \in \{\hat{s_L}, \hat{s_H}\}$, a draw from a signal structure we construct now, $\hat{\mathcal{P}}$. $\hat{x}$ is distributed independently from x conditional on θ, as follows:

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

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

$$\frac{\mathbb{P}\left((x, \hat{x}) = (s_L, \hat{s_H}) \mid \theta = H\right)}{\mathbb{P}\left((x, \hat{x}) = (s_L, \hat{s_H}) \mid \theta = H\right)} = \frac{s_H}{1 - s_H} \tag{4.5}$$

$$\frac{\mathbb{P}\left((x, \hat{x}) = (s_L, \hat{s_L}) \mid \theta = H\right)}{\mathbb{P}\left((x, \hat{x}) = (s_L, \hat{s_L}) \mid \theta = H\right)} = \frac{s_L}{1 - s_L} \times \frac{1 - \varepsilon \times \dfrac{s_H}{1 - s_H}}{1 - \varepsilon \times \dfrac{s_L}{1 - s_L}} \tag{4.6}$$

Note that 4.6 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 $(x, \hat{x})$ is informationally equivalent to $\mathcal{P}'$ when:

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

Choose ε to satisfy this equality; note that ε becomes a continuously increasing function of δ.

Step 2: $\pi(\sigma_1; \mathcal{P})$ increases (decreases) with a marginal decrease in s_L, whenever $s_L \geq s_L^{as}$.

The strategy σ_1 can be replicated by an evaluator who receives the signal pair $(x, \hat{x})$ instead of x'. To do so, the evaluator rejects if and only if the pair $(x, \hat{x}) = (s_L, \hat{s_L})$ is observed.

Fix the collection of signal draws evaluators will see for the applicant if he visits them all: $\{(x_i, \hat{x}_i)\}_{i=1}^n$. An applicant is a *marginal admit* if: (i) he has *no* $x = s_H$ signals, and (ii) he has *at least one* $\hat{x} = \hat{s_L}$ signal. The difference between evaluators' payoffs under $(\mathcal{P}, \hat{\mathcal{P}})$ and $\mathcal{P}$ is determined by these *marginal admits*, who are *eventually rejected* under $\mathcal{P}$, but *eventually approved* under $(\mathcal{P}, \hat{\mathcal{P}})$. So:

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

For a marginal admit, $(x_i, \hat{x}_i) \in \{(s_L, \hat{s_H}), (s_L, \hat{s_L})\}$, and $(x_j, \hat{x}_j) = (s_L, \hat{s_H})$ for at least one evaluator j. Denote the number of evaluators who observe $(s_L, \hat{s_H})$ as $\#$. Then, (2) equals:

$$\sum_{i=1}^{n} \underbrace{\frac{\mathbb{P}\left(i \ \hat{x} = \hat{s_H} \text{ signals } \mid x_1 = ... = x_n = s_L\right)}{\sum_{j=1}^{n} \mathbb{P}\left(j \ \hat{x} = \hat{s_H} \text{ signals } \mid x_1 = ... = x_n = s_L\right)}}_{(3)} \times \mathbb{P}\left(\theta = H \mid \# = i\right) - c$$

where:

$$\mathbb{P}\left(i \ \hat{x} = \hat{s_H} \text{ signals } \mid x_1 = ...x_n = s_L\right) = \binom{n}{i} \times (k \times \varepsilon)^i \times (1 - k \times \varepsilon)^{n-1}$$

for $k = \mathbb{P}\left(\theta = H \mid x_1 = ... = x_n = s_L\right)$. The limit of expression (3) as $\varepsilon \to 0$ (thus $\delta \to 0$) is:

$$\lim_{\varepsilon \to 0} \frac{\mathbb{P}\left(i \ \hat{x} = \hat{s_H} \text{ signals } \mid x_1 = ... = x_n = s_L\right)}{\sum_{j=1}^{n} \mathbb{P}\left(j \ \hat{x} = \hat{s_H} \text{ signals } \mid x_1 = ... = x_n = s_L\right)} = \mathbb{P}\left(1 \ \hat{x} = \hat{s_H} \text{ signals } \mid x_1 = ... = x_n = s_L\right)$$

and thus:

$$\lim_{\varepsilon \to 0} \mathbb{P}\left(\theta = H \mid \text{marginal admit}\right) - c = \lim_{\varepsilon \to 0} \mathbb{P}\left(\theta = H \mid \# = 1\right) - c$$

This expression is strictly positive (negative) when the expression below is strictly positive (negative):

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

there is a limit below as well, add.

proving the claim.

$\square$

Proposition 3. Let $\mathcal{P}$ be binary with normalised posterior beliefs $0 \leq s_L \leq 0.5$ and $0.5 \leq s_H \leq 1$. Evaluators' payoffs across the most embracive equilibria are:

i weakly increasing with the strength of evidence for $\theta = L$ when $s_L \geq \min\{s_L^{\text{mute}}, s_L^{as}\}$

ii weakly decreasing with the strength of evidence for $\theta = L$ when $s_L < \min\{s_L^{\text{mute}}, s_L^{as}\}$

Where $s_L^{\dagger} \geq \min\{s_L^{\text{mute}}, s_L^{as}\}$, their payoffs across the most selective equilibria are:

i weakly increasing with the strength of evidence for $\theta = L$ when $s_L \geq s_L^{\dagger}$

ii weakly decreasing with the strength of evidence for $\theta = L$ when $s_L < s_L^{\dagger}$

Proof. I start with the most embracive equilibrium. When $s_L \geq s_L^{\text{safe}}$, the strategy $\sigma_{(1,1)}$ which approves everyone; i.e. $\sigma_{(1,1)}(s_L) = \sigma_{(1,1)}(s_H) = 1$, is an equilibrium. This owes to $\Psi(\sigma_{(1,1)}; \mathcal{P}) = \rho$ as it can be easily checked, and to the definition of s_L^{safe}. Since no strategy is more embracive, $\sigma^{\text{em}*} = \sigma_{(1,1)}$. In this parameter region, $\pi(\sigma^{\text{em}*}; \mathcal{P})$ does not vary as every applicant is approved. When $s_L < s_L^{\text{safe}}$, this equilibrium is no longer possible, and evaluators' equilibrium payoffs are thus given by $\pi(\sigma^{\text{em}*}; \mathcal{P}) = \max\{0, \pi(\sigma_1; \mathcal{P})\}$; as it was explained in the second fact under Theorem 1's proof. As s_L decreases, this increases (decreases) when $s_L \geq s_L^{as}$ ($s_L < s_L^{as}$). This establishes the first part of Proposition 3.

For the most selective equilibrium to have $\sigma^*(\text{sel}*) = 0$, a necessary and sufficient condition is:

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

This owes to Lemma 4, which establishes that the interim belief *increases* in $\sigma(s_H)$.

Clearly, this condition is satisfied when $s_L \leq s_L^{\text{safe}}$. Thus, the most selective equilibrium becomes one where s_L leads to a rejection once s_L falls below some threshold $s_L^{\text{thr}} \geq s_L^{\text{safe}}$. Evaluators' equilibrium payoffs then start falling with stronger evidence for $\theta = L$ once $s_L \leq \min\{s_L^{\text{thr}}, s_L^{as}\}$.

□

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.

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

Muirhead, R. F. (1900). Inequalities relating to some algebraic means. *Proceedings of the Edinburgh Mathematical Society*, *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

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.