A COMPARISON OF NOISY SIGNALS IN SCREENING

David Lagziel and Ehud Lehrer

Discussion Paper No. 19-04

October 2019

Monaster Center for
Economic Research
Ben-Gurion University of the Negev
P.O. Box 653
Beer Sheva, Israel

Fax: 972-8-6472941
Tel: 972-8-6472286

A Comparison of Noisy Signals in Screening*

David Lagziel† and Ehud Lehrer‡

October 3, 2019

Abstract:

This paper studies the impact of different noises over screening processes. Specifically, it deals with a decision problem in which one decision maker screens a set of elements based on noisy unbiased evaluations. Given that the decision maker uses threshold strategies, we show that additional binary noise can potentially improve a screening - an effect that resembles a "lucky-coin toss". On the other hand, once optimal screening strategies are introduced, this effect disappears as any additional noise can only damage the screening process. We compare different noises under threshold strategies and optimal ones, and provide several (partial) characterizations of cases in which one noise is preferable over another. Doing so, we establish a novel method to compare noise variables through a mapping between percentiles.

Journal of Economic Literature classification numbers: C70, D49, D81.

Keywords: screening; threshold strategies; optimal screening.

*For their valuable comments, the authors wish to thank Omer Edhan, Dima Shaiderman, Eilon Solan, and Anna Zseleva, as well as the participants of ORSIS 2019 and the Tel-Aviv University Game Theory and Mathematical Economics Research Seminar. Lagziel acknowledges the support of the Israel Science Foundation, Grant #513/19.

†Ben-Gurion University of the Negev, Beer-Sheva 8410501, Israel. e-mail: Davidlag@bgu.ac.il.

‡Tel Aviv University, Tel Aviv 6997801, Israel and INSEAD, Bd. de Constance, 77305 Fontainebleau Cedex, France. e-mail: Lehrer@post.tau.ac.il.

1 Introduction

Imagine that you are on the verge of an important decision. To address this decision properly, you gather all relevant (potentially noisy) information and, based on that information, you dictate some decision rule, namely a strategy. Now assume you are approached by a person who advises you to strictly improve your decision, and that advice somewhat depends on a lucky coin. Hearing this, you will quite possibility consider the suggestion to be a joke, and for good reason. It just does not seem reasonable that one can improve a decision by introducing additional independent noise to the process. Nonetheless, in this paper we substantiate the potential superiority of the mentioned proposal by studying the impact of different noises on various screening problems.

This research begins with a screening problem in which one decision maker (DM) performs a screening based on noisy unbiased evaluations. The decision maker could be a manager reviewing job applicants, an editor of a peer-reviewed academic journal screening for insightful papers, or even a rating agency trying to asset the default risk of various borrowers. In all these scenarios (and various others), the decision problem is based on some noisy evaluation upon which the DM decides whether to accept or reject an element from a general set.

Following the model of Lagziel and Lehrer (2019), we assume that the accurate values of the elements in question are distributed according to an impact variable V and there exists a noise variable N, such that the DM observes $V + N$ while trying to maximize the expected impact of accepted elements through a proper decision rule (i.e., a screening strategy) which depends on $V + N$. To ensure non-trivial results, we assume that the DM has a capacity constraint such that a certain volume of elements must eventually be accepted.

The search for optimal screening strategies typically begins by examining threshold strategies for an obvious reason - in the absence of noise, threshold strategies are indeed optimal. In this paper, we adhere to this line of thinking as well. Indeed, our first main result roughly states that, under threshold strategies, one can strictly improve a screening by adding independent binary noise to the evaluations. In other words, we establish the possibility to generate "lucky coins" that improve a screening process. To provide some intuition for this statement, we highlight two key conditions that are essential for the mentioned result. The first is that threshold strategies be applied although they are not necessarily optimal. The second is that the original noise N can generate non-trivial ordinal changes among values of V, conditional on $V + N$. Once some ordinal changes occur, the additional noise can partially correct the applied screening strategy. A concrete (and rather simple) example of this result is given in Section 3.1.

This preliminary result is merely the overture to a much broader question concerning the way different noises impact screening processes; a question that stands at the core of the current work. We address this research question by examining the superiority of one noise variable over another. That is, we ask whether, *ceteris paribus*, a screening under one noise variable produces a better result than under a different one. More formally, we refer to this situation as *screening dominance*, and say that one noise variable *S-dominates* another if the expected value of accepted elements given the former noise is at least as high as the expected value given the latter.

Our second main result provides a partial characterization of screening dominance under threshold strategies. Specifically, we fix two non-atomic noises and define a percentile-transformation (PT) function between the two noises. Our equivalence result shows that one noise S-dominates another if and only if the PT function is a contracting mapping. This result establishes a new method to compare noise variables, namely a contraction mapping, which defers from commonly known ones such as the mean-preserving spread (see literature review below for more details).

The next stage of our analysis focuses on optimal screening strategies. We begin this stage by proving that the lucky-coin outcome completely changes once optimal strategies are introduced. Assuming that optimal screening strategies are applied, we prove that a lucky coin cannot exist since additional noise can only damage the screening process. This result leads to a characterization of screening dominance between normally distributed noises. That is, we consider two normally distributed noises N_1 and N_2, and prove that N_1 S-dominates N_2 if and only if N_1 could be generated by the sum of N_2 and another normally distributed, independent noise.

The last part of our analysis combines the previously mentioned results by proving that threshold strategies are optimal once uniform non-atomic noises are considered. Given such noises, we show that screening dominance is not fully characterized by additive noise, but may follow from the same contraction property that provided a key characterization under threshold strategies.

1.1 Related literature and main contribution

The economic research of screening and noisy signals ranges from job-market signalling and education to insurance and credit markets (see, e.g., Spence (1973), Stiglitz (1975), Rothschild and Stiglitz (1976), and Stiglitz and Weiss (1981)). Note that these papers (among many others) typically focus on costly screening and strategic signalling, while we consider a non-strategic and costless signalling model. So, in the relevant literature, the papers that are closest to ours are Rothschild and Stiglitz (1970, 1971), and Lagziel and Lehrer (2019).

The first aspect that associates our work with that of Rothschild and Stiglitz (1970, 1971) is the underline goal: relating probabilistic properties of random variables to the preferences of a rational decision maker. Rothschild and Stiglitz (1970, 1971) achieve this goal by defining the notion of a mean-preserving spread (MPS) which induces a partial order over lotteries, and then relating this order to the preferences of a risk-averse expected-utility maximizer. In contrast, in our paper, we consider an additive independent noise and provide several equivalence results between the induced partial order (over noises) with screening dominance.

This first similarity naturally leads to the second important connection between the studies - the origin of the partial orders in question. We, similarly to Rothschild and Stiglitz (1970, 1971), use additive independent noise as the basis for our partial order and analysis. However, Rothschild and Stiglitz use this noise to define the notion of a MPS, whereas we use it to define a contraction between noises which entails superior screening capabilities, either through threshold strategies or through optimal ones. In addition, we provide a combination of positive and negative results, specifically because we do not confine ourselves to optimal screening, but allow for commonly-used threshold strategies.

Another main resemblance between the studies is the ability to provide a wide range of applications for the given theoretical results. Rothschild and Stiglitz (1971) apply their earlier results (from the 1970 paper) to various investment and production problems. We, however, follow the model of Lagziel and Lehrer (2019) with its broad set of applications that range from peer-reviewed academic publishing to credit ratings.

Lastly, we wish to emphasize the strong connection between the current study and Lagziel and Lehrer (2019). Not only do we use a similar screening model, but the basic notion of a *screening bias*, which lies at the core of Lagziel and Lehrer (2019), is the starting point of the current work. We establish our lucky-coin result by first constructing a screening bias, and then partially correcting it using additive noise (e.g., see the figures and example given in Section 3.1). Note that the current work provides a much broader set of positive results and characterizations, specifically because we account for optimal strategies as well as threshold ones, while Lagziel and Lehrer (2019) only account for the latter.

1.2 Structure of the paper

The paper is organized as follows. In Section 2 we present the basic screening model. In Section 3 we study screening problems under threshold strategies and, in Section 4, we focus on screening problems under optimal strategies. In Section 4.2, we combine the results of Sections 3 and 4 by analysing screening problems with uniform noises. Concluding remarks are given in Section 5.

2 Preliminaries

We follow a basic screening model with one decision maker (DM) who performs a screening. Consider a set of elements whose values are distributed according to a non-constant random variable V, referred to as an *impact variable*. The elements' individual values are private, so every element with private value v goes through a noisy evaluation process and is evaluated by $v + N$, where N is an *unbiased* noise variable, i.e., it is symmetrically distributed around zero and independent of V.[1]

The DM uses the noisy evaluation to perform a screening subject to a capacity constraint. For this purpose, the DM sets a screening strategy $\sigma : \mathbb{R} \to \{0, 1\}$ where 1 denotes the acceptance of a specific valuation and 0 denotes a rejection.[2] To avoid trivial solutions, we fix a minimal rate of acceptance, a *capacity* level $p \in (0, 1)$, such that the share of accepted elements does not fall below p, and every screening strategy σ must ensure that $\Pr(\sigma(V + N) = 1) \geqslant p$.

To motivate this model, Lagziel and Lehrer (2019) provide an example in which the DM is an editor of a peer-reviewed academic journal who approaches referees to evaluate a set of academic papers: V denotes the papers' potential impact, $V + N$ is the referees' evaluations, and σ is the editor's decision rule to either accept or reject a paper. Other possible scenarios scenarios include a trader facing different investment opportunities, a manager screening potential employees, or even a sports scout searching for potential Hall-of-Fame players. In all these scenarios, the DM establishes a noisy screening process in order to maximize the expected value of accepted elements, subject to some capacity constraint.

We generally refer to the tuple $\mathrm{SP} = (V, N, p)$ as a *screening problem*. Given a screening problem SP and a screening strategy σ, the expected value of accepted elements is

$$\Pi_{\mathrm{SP}}(\sigma) \quad = \quad \mathbb{E}\left[V | \sigma(V + N) = 1\right].$$

The DM's goal is to maximize Π_{SP}. We denote the DM's optimal screening strategy and optimal expected payoff by σ^*_{SP} and Π^*_{SP}, respectively. To be clear, all definitions and statements hold almost surely (i.e., hold up to a measure-zero deviation).

The search for optimal screening typically begins by analysing the class of threshold strategies for two main reasons. The first is that, in the absence of noise, threshold strategies are optimal. The second is that, given a capacity p, threshold strategies are rather simple to implement since they are characterized by a unique threshold value which captures the top $100p$ percentile of the distribution. Thus, we will devote a portion of our analysis to study threshold strategies, and the non-trivial cases in which they are optimal.

Formally, a screening strategy σ is a *threshold strategy* if there exists a value s such that, with probability one, every noisy valuation (i.e., signal) above s is accepted and every noisy valuation below s is rejected. Given a screening problem SP, we denote a threshold strategy and the expected payoff under a threshold strategy by $\hat{\sigma}_{\mathrm{SP}}$ and $\widehat{\Pi}_{\mathrm{SP}}$, respectively.

Since we incorporate general distributions in this model, one final clarification is needed for the case of atomic ones. Should $V + N$ have an atomic distribution and to meet the capacity constraint p,

[1]Throughout this paper and unless stated otherwise, the notations N and N_i refer to unbiased noise variables.

[2]We henceforth assume that all measurablity requirements hold.

the DM may need to impose a partially random screening such that valuations which are subject to an atom, are randomly split. In such cases, one should consider a more general screening strategy where $\sigma : \mathbb{R} \to \Delta(\{0,1\})$. We typically abstract from these cases by assuming that (through an appropriate randomization) the DM can "split the atom" (in a mathematical sense) and capture the expected value given that atom, with the needed proportion.

2.1 Screening dominance and noisy amplifications

There are two noise-related notions that govern our analysis: screening dominance and noisy amplifications. Let us define and explain each of these notions, starting with the former.

Definition 1. [Screening dominance]. *Noise variable* N_1 S-dominates *noise variable* N_2 *if, for every impact variable* V *and capacity* p*, an optimal screening in* $\mathrm{SP}_1 = (V, N_1, p)$ *produces a higher expected value than an optimal screening given* $\mathrm{SP}_2 = (V, N_2, p)$*. That is,* N_1 S-dominates N_2 *if*

$$\Pi^*_{\mathrm{SP}_1} \geqslant \Pi^*_{\mathrm{SP}_2}$$

and the inequality is strict for some impact variable and capacity.

In simple terms, a noise variable N_1 S-dominates N_2 if, *ceteris paribus*, an optimal screening under the former noise is at least as good as an optimal screening under the latter (and, in some cases, strictly better), independently of either the impact or the capacity.

The notion of screening dominance is rather demanding in the sense that it requires superiority for every impact variable and every capacity under optimal strategies. In some cases we shall require a weaker notion such that optimal screening strategies are replaced with threshold ones. For such purposes, we say that N_1 *S-dominates* N_2 *under threshold strategies* if $\widehat{\Pi}_{(V,N_1,p)} \geqslant \widehat{\Pi}_{(V,N_2,p)}$ for every (V, p), and the inequality is strict for some impact variable and capacity.

The second notion we shall use is termed *noisy amplification*, and it suggests that one noise could be reproduced by another, through an independent lottery.

Definition 2. [Noisy amplification]. *A noise variable* N_2 *is a* noisy amplification *of noise* N_1 *if* $N_2 \sim N_1 + N_3$*, and noise variable* N_3 *is independent of* N_1*.*

In other words, N_2 is a noisy amplification of N_1 if one can produce the distribution of N_2 using the sum of N_1 and an independent lottery N_3. This notion will prove useful when debating the existence of lucky coins and the dominance of one noise over another.

3 Screening under threshold strategies

This section is divided into two parts, each part presenting one key result of the paper. The first part, given in Section 3.1, discusses the possibility to generate lucky coins - a simple lottery that strictly improves a screening. The second part, given in Section 3.2, presents the first partial characterization of screening dominance. In both parts we restrict our attention to threshold strategies that will be later combined, in Section 4.2, with the optimal ones.

3.1 A lucky coin toss for screening

The concept of a lucky coin toss is ambivalent. On the one hand, the procedure itself is simple, not to say trivial: Once a DM approaches some screening problem, she can simply toss a coin and incorporate its result into her decision. On the other hand, how can a simple coin toss improve a screening if we are merely introducing random noise to the screening process? In this section, we shall attempt to resolve this enigma.

We begin with a straightforward result stating that lucky coins exist. Theorem 1 below shows that for every bounded impact variable V and every capacity p, one can devise a noise variable N_1 such that a lucky coin exists for the screening problem SP $= (V, N_1, p)$. The introduction of a lucky coin toss is manifested through a different noise variable, N_2, which is a noisy amplification of N_1.

Theorem 1. *For every bounded impact variable V and capacity p, there exist noise variables N_1 and N_2 where N_2 is a noisy amplification of N_1, and $\widehat{\Pi}_{(V,N_2,p)} > \widehat{\Pi}_{(V,N_1,p)}$*

The implications of Theorem 1 are clear: In some cases one can strictly improve a screening by inserting additional noise to the process. To clarify the last statement and explain our use of the lucky-coin terminology, we remark that the proof of Theorem 1 uses an amplification of N_1, namely $N_2 \sim N_1 + N_3$, where N_3 is a *binary* symmetric noise.

We now turn to focus on the driving force behind this result. First, recall the key insight of Lagziel and Lehrer (2019) stating that $\widehat{\Pi}_{(V,N,p)}$ is not necessarily a monotonic function of p. In other words, the DM can enforce a more restrictive screening and the expected average level can actually decrease. The non-monotonicity of $\widehat{\Pi}_{(V,N,p)}$ w.r.t. p follows from the fact that unbiased noise has a different nominal effect over different-size sets to the point that it significantly distorts the impact variable's conditional distribution. That is, an unbiased noise imposed over a large set of mediocre elements, will produce a non-negligible amount of upwards shifting, whereas the same noise imposed over a small set of superior elements produces a relatively small amount of upwards shifting. In such cases, the probability masses matter, and unbiased noise can distort the distribution (generating, what is called, "a screening bias"), such that the noisy valuations do not reflect the "true" ordering of the elements' impact.

The lucky coin toss can partially deal with this problem, as shown in the simplified example given in Figures 1 and 2, motivated by the main result of Lagziel and Lehrer (2019). In Figure 1, the DM receives a distribution of noisy valuations where the values in parenthesis denote the true (i.e., accurate) assessment subject to the given signal. If the capacity is fixed at $p = 0.5$ and the DM is using a threshold strategy, then 18% of A elements and 32% of $B+$ elements would be accepted (note that a noisy signal of $A-$, actually reflects a true value of $B+$). However, once a simple coin toss is introduced such that every valuation shifts either upward or downward by one level with equal probabilities, then the DM observes the distribution that is given in Figure 2. Holding the capacity fixed, the DM would now accept 18% of A elements, 16% of $A-$ elements, and only 16% of B elements. The lucky coin shifted the signals in such a way that a portion of $B+$ elements are replaced by the same mass of $A-$ elements, thus increasing the expected value of accepted elements.

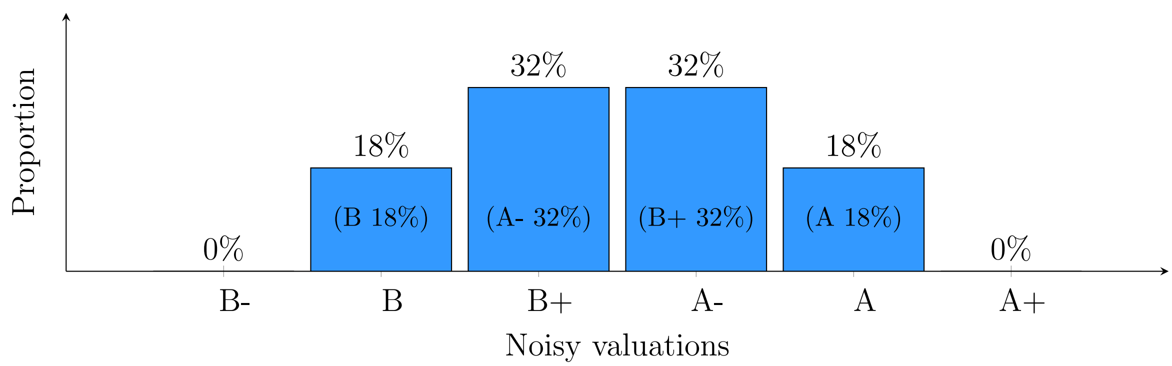

Figure 1: The noisy evaluation *prior* to the lucky coin. Values in parenthesis reflect the accurate valuations.

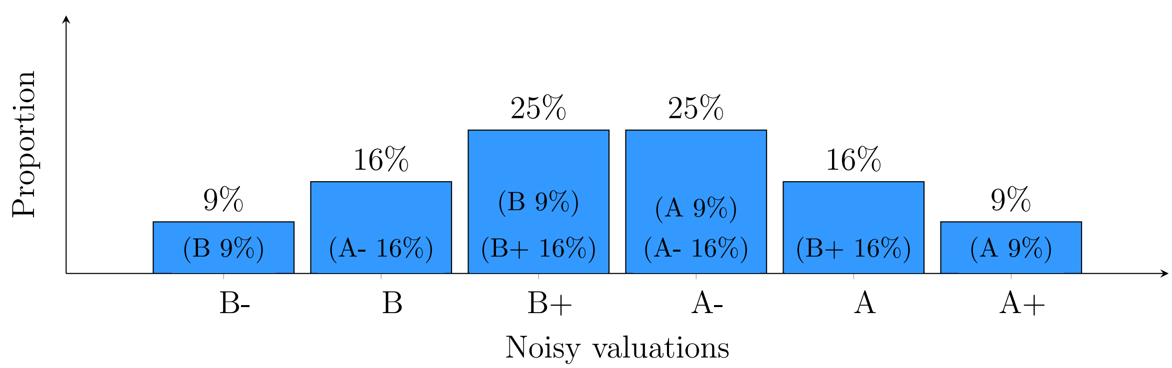

Figure 2: The noisy evaluation *after* the lucky coin is introduced. Values in parenthesis reflect the accurate valuations.

An important component of this result is that we are applying threshold strategies. Should the DM have the ability to apply optimal strategies, then the same example would show how the additional noise only damages the screening process. We will return to this issue in Section 4 when discussing screening under optimal strategies.

3.1.1 The robustness of the lucky coin toss

We wish to discuss two robustness concerns regarding the result of Theorem 1 and, specifically, the relevant noises for that result.

First, the noises used in the proof of Theorem 1 depend only on whether $p \geqslant 0.5$ or $p < 0.5$, and on the support of V rather than on its entire distribution. In addition, the dependency on p hinges on the need to sustain symmetric noises. Therefore, if one should allow for asymmetric noises, then the result of Theorem 1 becomes rather general, in that one can generate a strictly better screening for every impact variable V (with the same support) and for every capacity p, while holding N_1 and N_3 fixed.

Second, the support of the additive noise used in the proof is quite narrow relative to the support of V; this is also evident from the previous example, given in Figures 1 and 2. Therefore, the lucky coin can improve the screening although its magnitude, in general, is rather small. The fact that the ordinal changes are locally generated and the use of threshold strategies suggest that even if additional valuations are introduced (enlarging the support of V), the result of Theorem 1 would still hold.

3.2 A partial characterization of screening dominance

To characterize screening dominance under threshold strategies, we begin by defining a function which transforms any non-atomic noise to another non-atomic noise using percentiles translation. This function will be used to transform a threshold strategy under one noise variable to a different threshold strategy under another noise variable. As it turns out, the key property to determine whether one noise variable S-dominates another is whether this function is a contraction mapping or not. If the function is a contraction, meaning that one noise transforms to another using some form of contraction, then the condensed noise is superior for screening purposes.

Formally, consider two noise variables N_1 and N_2 with CDFs F_1 and F_2, respectively. For the sake of simplicity, assume both noises are non-atomic with convex supports such that percentiles are uniquely defined. Given such noises, define the *Percentile-Translation* (PT) function T_{ij} by

$$T_{ij}(n) = F_i^{-1}\left(F_j(n)\right), \quad \forall n \in \mathrm{Supp}(N_j).$$

In other words, the PT function receives as input any $100p$-percentile of noise N_j and generates the $100p$-percentile of N_i.

Let us now review the key properties of the PT function. Since both noises are non-atomic with convex supports, the CDFs are strictly increasing on these sets and the function $T_{ij}(n)$ is well-defined and strictly increasing, as well. Second, it is straightforward to verify that T_{12} is the inverse of T_{21} and both are bijective functions (one-to-one correspondences) between the relevant supports. In that case, $T_{21} = T_{12}^{-1}$ as well as $T'_{21}(n) = \frac{1}{T'_{12}(T_{21}(n))}$. To simplify the exposition, we apply two additional definitions: (i) N_i and N_j are called *continuous* if both noises are non-atomic with convex supports, while T_{ij} and T_{ji} are continuously differentiable; and (ii) N_i is called a *contraction* of N_j if $T'_{ij}(n) \leqslant 1$ for every $n \in \mathrm{Supp}(N_j)$ while T_{ij} is continuously differentiable.

As follows from Theorem 2 below, the contraction property of the TP function is a necessary and sufficient condition for screening dominance under threshold strategies. Our equivalence result states that N_1 S-dominates N_2 under threshold strategies if and only if the TP mapping T_{12} is a contraction.

Theorem 2. *Fix two distinct continuous noise variables N_1 and N_2. Then, N_1 is a contraction of N_2 if and only if N_1 S-dominates N_2 under threshold strategies.*

Figure 3 provides intuition for the proof of Theorem 2. The graph on the left represents threshold-screening under N_2 where l_2 denotes the threshold line for that screening. Using the PT function, one can translate l_2 to terms of N_1, thus getting the line l'_2 on the RHS graph. Note that the transformation along with the contraction property of T_{12} ensure that l'_2 is decreasing with a slope greater than -1. When comparing the screening according to l'_2 with the threshold-screening w.r.t. N_1 (given by the grey areas on the RHS and l_1 line), we see that lower values of V are discarded in favor of higher ones (light-grey area B instead of area A). Hence, the threshold-screening under N_1 is superior as stated.

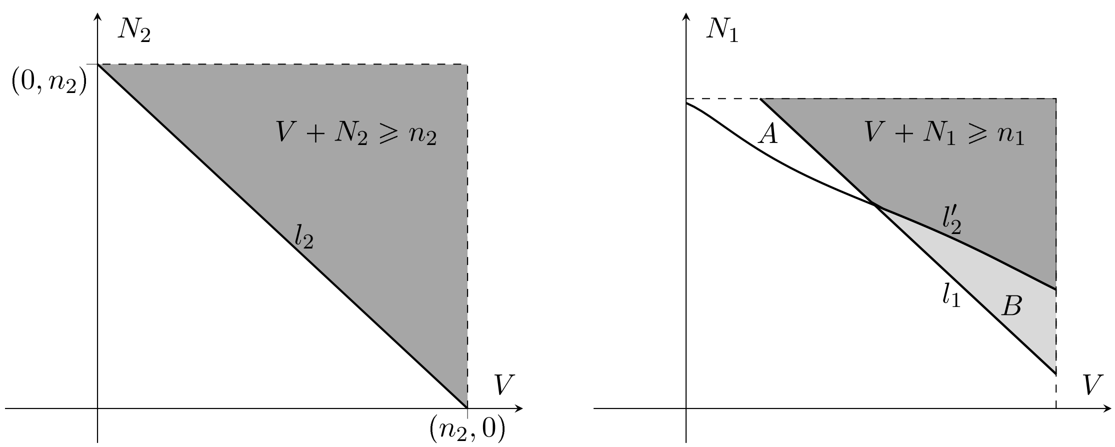

Figure 3: Each graph $l_i : V + N_i = n_i$ represents threshold-screening in screening problem SP_i, and each shaded area $V + N_i \geqslant n_i$ represents the accpeted valuations. Line $l'_2 : N_1 = T_{12}(n_2 - V)$ is the representation of l_2 in terms of N_1 using the PT function T_{12}. This translation maintains capacity, and its slope is greater than -1. Given threshold-screening under N_1, set A (white area) is replaced by set B (light grey area), ensuring the screening dominance of N_1.

What happens if the PT function is not a contraction? Assuming that noises are distinct and unbiased, then T_{12} is not a linear function. Thus, there exists a point n such that $T'_{12}(n) > 1$, which suggests that T_{12} is locally expanding in the neighbourhood of n (recall that the PT function is continuously differentiable) and T_{21} is locally contracting on some interval. Hence, one can choose an arbitrary small-support impact variable such that the local contraction of T_{21} generates the same effect as shown in Figure 3, but when translating N_1 to N_2. This guarantees that the existence of an impact variable and capacity such that threshold-screening under N_2 is superior (for a detailed proof, see Lemma 4 in the Appendix).

A simple example of a contracting PT function follows from multiplying a noise variable by any positive constant $c \in (0, 1)$. Once this is done, the resulting noise is a contraction of the former. We shall return to this simple observation when discussing optimal screening under uniform noises in Section 4.2.

We conclude this section with an explanation concerning the requirement for continuous noises. Consider, for example, two binary and symmetric noise variables N_1 and N_2 where $N_i = \pm i$ with equal probabilities. Now take an impact variable V which equals either 0 or 2, again, with equal probabilities. Under the N_1 noise, the DM would get a signal $s = 1$, but she would not know whether it originated from the combination $(V, N_1) = (2, -1)$ or from $(V, N_1) = (0, 1)$. In contrast, such ambiguity does not occur under noise N_2 which would generate four distinct signals $s \in \{-2, 0, 2, 4\}$. This type of information ambiguity implies that N_1 does not S-dominate N_2, and exemplifies the necessity of our continuity requirement.

4 Screening under optimal strategies

In this section we focus on screening dominance under optimal strategies, and for that purpose we divide our analysis into two parts. First, in Section 4.1, we show that a noisy amplification is a sufficient condition for screening dominance. Moreover, if one should restrict attention to normally distributed noises, we prove that the noisy amplification condition is, in fact, a characterization of screening dominance. Second, in Section 4.2, we examine the noisy amplification condition under uniform noises. Specifically, we prove that a noisy amplification is not a necessary condition for screening dominance under uniform noises, whereas the contraction result of Theorem 2 does provide a necessary condition under such noises.

4.1 A sufficient condition for screening dominance

The first result connects the two basic notions of noisy amplifications and S-dominance. Specifically, Theorem 3 below states that a noisy amplification of one noise variable is dominated, in terms of screening, by that variable.

Theorem 3. *If N_2 is a noisy amplification of N_1, then N_1 S-dominates N_2.*

In order to prove Theorem 3, one needs to devise an optimal screening strategy for general screening problems. Thus, an important insight that one can extract from the proof of Theorem 3 is the structure of that optimal strategy. Namely, the optimal strategy used in this proof is based on the function

$$f_i(s) = \mathbb{E}[V|V + N_i = s],$$

for every signal s. In words, the function provides the expected value of the impact variable conditional on the received signal. For every capacity p, the optimal strategy dictates to accept a noisy valuation of s if $f_i(s) \geqslant t_{\mathrm{SP}_i}$ for some fixed value t_{SP_i}, which depends on the screening problem SP_i. This carries some resemblance to the Neyman–Pearson lemma, and one can also find some similarities between the two proofs.

An immediate conclusion from Theorem 3 is an equivalence between screening dominance and noisy amplifications within the set of normally distributed noises. The driving force behind this conclusion is the fact that the set of normally distributed unbiased noises is closed with respect to additivity, and that for any two such distinct noises N_1 and N_2, either N_1 is a noisy amplification of N_2 or vice versa.

Corollary 1. *Fix two normally distributed noise variables N_1 and N_2. Then, N_2 is a noisy amplification of N_1 if and only if N_1 S-dominates N_2.*

The proof is straightforward (and thus omitted). One direction follows directly from Theorem 3, so we should only consider the other direction starting with the screening dominance of N_1 over N_2. If N_1 S-dominates N_2, there exists an impact variable and capacity such that screening under N_1 is strictly better. Thus, the two noises are not distributed similarly and one has a higher variance than the other. The noise with the higher variance is a noisy amplification of the other, and it is evident (from Theorem 3) that N_2 is a noisy amplification of N_1.

4.2 Screening dominance under uniform noises

Corollary 1 naturally gives rise to a follow-up question: Is there an equivalence between screening dominance and noisy amplifications under general distributions? It appears that the answer to this question is negative, since one cannot identify screening dominance by solely restricting attention to noisy amplifications. We show this by focusing on the class of uniformly distributed convex-support noises.

We begin our analysis by establishing, in Lemma 1 below, that the optimal screening strategies under uniformly distributed convex-support noises are threshold strategies.

Lemma 1. *If N is uniformly distributed on an interval, then threshold strategies are optimal for every V and p.*

Given this result, we can consider a simple transformation of noises, other than additive noise, that damages the screening process. Specifically, we can multiply a noise variable by a constant greater than one, and analyse how the expansion affects the screening. In other words, we can fix two continuous noise variables N_1 and N_2 (as considered in Section 3.2) where $N_2 \sim cN_1$ and $c > 1$. It is easy to verify that N_1 is a contraction of N_2, thus the former S-dominates the latter under (the optimal) threshold strategies. In the following lemma, we prove this result for the general distributions without confining ourselves to continuous noises.

Lemma 2. *Fix two screening problems* $\mathrm{SP}_i = (V, N_i, p)$ *where* $i = 1, 2$ *and* $N_2 \sim cN_1$ *for some* $c > 1$. *Then,* $\widehat{\Pi}_{(V,N_1,p)} \geqslant \widehat{\Pi}_{(V,N_2,p)}$.

Note that the statement of Lemma 2 is general, and independent on the distribution of N_1. In other words, this result is not limited to either uniform, or continuous noises.

Using Lemma 1, Lemma 2 and Lemma 3 below, we can prove that noisy amplifications do not provide a general characterization for dominance. Specifically, fix two uniformly distributed[3] noises $N_1 \sim U[0, 1]$ and $N_2 \sim U[0, 1.5]$. Lemma 1 states that the optimal screening strategy in any screening problem (under these noises) is a threshold one. Lemma 2 establishes that N_1 dominates N_2 since $N_2 \sim 1.5N_1$. Lemma 3, which follows, proves that N_2 is not a noisy amplification of N_1. Thus, we substantiate the existence of two noise variables such that one noise S-dominates another, while the noisy-amplification condition is violated.

Lemma 3. *If* $N_1 \sim U[0, 1]$ *and* $N_2 \sim U[0, 1.5]$, *then* N_2 *is not a noisy amplification of* N_1.

In other words, one cannot devise an independent noise N such that $N_2 \sim N_1 + N$, nonetheless, N_1 S-dominates N_2.

5 Conclusion

In this paper we provided several novel insights to the world of screening. Using our definition of *screening dominance*, we showed that additional noise is not necessarily adversary for a DM, assuming that threshold strategies are exercised. Next, we compared various noises in the context of screening while accounting for threshold strategies as well as optimal ones. We were able to provide several characterizations for screening dominance among different type noises, and most importantly, our main characterization result shows that some form of contraction among the noises' distributions is essential for screening dominance.

References

Lagziel, D. and E. Lehrer (2019): "A Bias of Screening," *American Economic Review: Insights.*

[3] For the sake of simplicity, we do not consider symmetric noises. However, modifying the example to symmetric noises is straightforward.

Rothschild, M. and J. Stiglitz (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 90, 629.

Rothschild, M. and J. E. Stiglitz (1970): "Increasing risk: I. A definition," *Journal of Economic Theory*, 2, 225–243.

——— (1971): "Increasing risk II: Its economic consequences," *Journal of Economic Theory*, 3, 66–84.

Spence, M. (1973): "Job Market Signaling," *The Quarterly Journal of Economics*, 87, 355–374.

Stiglitz, J. E. (1975): "The Theory of "Screening," Education, and the Distribution of Income," *American Economic Review*, 65, 283–300.

Stiglitz, J. E. and A. Weiss (1981): "Credit Rationing in Markets with Imperfect Information," *The American Economic Review*, 71, 393–410.

A Appendices

A.1 Proof of Theorem 1

Proof. Fix an impact variable V and a capacity $p \in (0,1)$. Assume, w.l.o.g., that V is supported on $[0,1]$. We separately relate to three cases: $p < 0.5$, $p = 0.5$, and $p > 0.5$. In general, denote the screening problem $\mathrm{SP}_i = (V, N_i, p)$ for every i and every noise N_i.

Starting with $p < 0.5$, define the noise variable N_1 by

$$N_1 = \begin{cases} \pm 1.1, & \text{w.p. } p, \\ 0, & \text{w.p. } 1-2p. \end{cases}$$

Evidently, $\mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_1}(V+N_1) = 1] = \mathbb{E}[V]$. Now consider $N_3 = \pm 0.1$ with equal probabilities and $N_2 \sim N_1 + N_3$. The distribution of N_2 is therefore

$$\Pr(N_2 = k) = \begin{cases} p/2, & \text{for } k \in \{\pm 1.2, \pm 1\}, \\ 1/2 - p, & \text{for } k \in \{\pm 0.1\}. \end{cases}$$

Given $V + N_2$, the threshold strategy $\hat{\sigma}_{\mathrm{SP}_2}$ accepts every assessment once $N_2 = 1.2$ and only partially accepts assessments once $N_2 = 1$. The latter is due to the fact that, given $N_2 = 0.1$, some high values of V are accepted instead of low values of V, given $N_2 = 1$. We conclude that the threshold level is some $t \in [1, 1.1]$. Thus,

$$\begin{aligned}
\mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_2}(V+N_2)=1] &= \frac{\mathbb{E}\left[V\mathbf{1}_{\{V+N_2\geqslant t\}}\right]}{p} \\
&= \frac{\frac{p}{2}\mathbb{E}\left[V\mathbf{1}_{\{V+1.2\geqslant t\}}\right]+\frac{p}{2}\mathbb{E}\left[V\mathbf{1}_{\{V+1\geqslant t\}}\right]+\left(\frac{1}{2}-p\right)\mathbb{E}\left[V\mathbf{1}_{\{V+0.1\geqslant t\}}\right]}{p} \\
&= \frac{\mathbb{E}[V]}{2}+\frac{1}{2}\mathbb{E}\left[V\mathbf{1}_{\{V\geqslant t-1\}}\right]+\left(\frac{1}{2p}-1\right)\mathbb{E}\left[V\mathbf{1}_{\{V\geqslant t-0.1\}}\right] \\
&= \frac{\mathbb{E}[V]}{2}+\frac{\Pr(V\geqslant t-1)}{2}\mathbb{E}[V|V\geqslant t-1] \\
&+ \Pr(V\geqslant t-0.1)\left(\frac{1}{2p}-1\right)\mathbb{E}[V|V\geqslant t-0.1] \\
&> \mathbb{E}[V]\left[\frac{1}{2}+\frac{\Pr(V\geqslant t-1)}{2}+\Pr(V\geqslant t-0.1)\left(\frac{1}{2p}-1\right)\right] \\
&= \mathbb{E}[V]\frac{\Pr(V+N_2\geqslant t)}{p}=\mathbb{E}[V].
\end{aligned}$$

Therefore, $\mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_2}(V+N_2)=1] > \mathbb{E}[V] = \mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_1}(V+N_1)=1]$.

For the case of $p > 0.5$, follow a similar computation while substituting p with $1-p$. This will produce the same inequality $\mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_2}(V+N_2)=1] > \mathbb{E}[V] = \mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_1}(V+N_1)=1]$.

For the case of $p = 0.5$, define $N_1 = \pm 0.6$ with equal probabilities, and define $N_3 = \pm 0.2$ with equal probabilities, as well. Hence, $N_2 \in \{\pm 0.8, \pm 0.4\}$ all with equal probabilities. Clearly, $\mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_1}(V+N_1)=1] = \mathbb{E}[V]$, while the screening threshold under N_2 is some value $t_2 \in [0.4, 0.6]$, and

$$\begin{aligned}
\mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_2}(V+N_2)=1] &= \mathbb{E}[V|V+N_2\geqslant t_2] \\
&= \frac{1}{4}\mathbb{E}\left[V\mathbf{1}_{\{V\geqslant t_2-0.8\}}\right]+\frac{1}{4}\mathbb{E}\left[V\mathbf{1}_{\{V\geqslant t_2-0.4\}}\right]+\frac{1}{4}\mathbb{E}\left[V\mathbf{1}_{\{V\geqslant t_2+0.4\}}\right] \\
&= \frac{1}{4}\mathbb{E}[V]+\frac{\Pr(V\geqslant t_2-0.4)}{4}\mathbb{E}[V|V\geqslant t_2-0.4] \\
&+ \frac{\Pr(V\geqslant t_2+0.4)}{4}\mathbb{E}[V|V>t_2+0.4] \\
&> \mathbb{E}[V]\left[\frac{1}{4}+\frac{\Pr(V\geqslant t_2-0.4)}{4}+\frac{\Pr(V\geqslant t_2+0.4)}{4}\right] \\
&= \mathbb{E}[V]=\mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_1}(V+N_1)=1],
\end{aligned}$$

which concludes the proof. ■

A.2 Proof of Theorem 2

For the proof of Theorem 2 we require the following auxiliary lemma.

Lemma 4. *Consider two continuous noise variables N_1 and N_2. For every $n \in \mathrm{Supp}(N_2)$ such that $T'_{12}(n) < 1$, there exists (V, p) such that $\widehat{\Pi}_{(V,N_1,p)} > \widehat{\Pi}_{(V,N_2,p)}$. Moreover, if T_{12} is a contraction, then $\widehat{\Pi}_{(V,N_1,p)} \geqslant \widehat{\Pi}_{(V,N_2,p)}$ for every (V, p).*

Proof. Fix two distinct continuous noise variables $\{N_1, N_2\}$, and take an interior point $n_2 \in \mathrm{Supp}(N_2)$ such that $T'_{12}(n_2) < 1$. Since T_{12} is continuously differentiable, one can take an open interval $I = (n_2 - \varepsilon, n_2 + \varepsilon)$ and get $T'_{12}(n) < 1$ for every $n \in I$. Define $V \sim U[-\varepsilon, \varepsilon]$, and consider the screening problem $\mathrm{SP}_2 = (V, N_2, p)$ where p is fixed such that $\hat{\sigma}_{\mathrm{SP}_2}(s) = 1$ if and only if $s \geqslant n_2$. That is, the threshold-screening for (V, N_2, p) accepts every valuation given by the event $\{V + N_2 \geqslant n_2\}$.

Note that $N_1 \sim T_{12}(N_2)$ since, for every $n \in \mathbb{R}$, we get

$$\Pr(T_{12}(N_2) \leqslant n) = \Pr(F_1^{-1}(F_2(N_2)) \leqslant n) = \Pr(N_2 \leqslant F_2^{-1}(F_1(n))) = F_2(F_2^{-1}(F_1(n))) = F_1(n).$$

So, T_{12} transforms N_2 to N_1. Hence,

$$\begin{aligned}
\mathbb{E}[V|V + N_2 \geqslant n_2] &= \mathbb{E}[V|N_2 \geqslant n_2 - V] \\
&= \mathbb{E}[V|T_{12}(N_2) \geqslant T_{12}(n_2 - V)] \\
&= \mathbb{E}[V|N_1 \geqslant T_{12}(n_2 - V)],
\end{aligned}$$

where the second equality holds by the fact that T_{12} is strictly increasing.

Let us consider the function $f(v) = T_{12}(n_2 - v)$ for $v \in (-\varepsilon, \varepsilon)$. Clearly, this function is strictly decreasing and differentiable such that $f'(v) = -T'_{12}(n_2 - v) > -1$ for every $v \in (-\varepsilon, \varepsilon)$. For every $c \in (-\varepsilon, \varepsilon)$, define the linear function $g_c(v) = -v + c + T_{12}(n_2 - c)$. Note that $g'_c(v) = -1$, so the functions $f(v)$ and $g_c(v)$ intersect exactly once at $(c, T_{12}(n_2 - c))$. Specifically, $g_\varepsilon(v) \geqslant f(v)$ whereas $g_{-\varepsilon}(v) \leqslant f(v)$.

We can now use g_c to construct a threshold (screening) strategy for screening problem (V, N_1, p). Observe that

$$\Pr(N_1 \geqslant g_\varepsilon(V)) < \Pr(N_1 \geqslant f(V)) = \Pr(N_1 \geqslant T_{12}(n_2 - V)) = p,$$

while

$$\Pr(N_1 \geqslant g_{-\varepsilon}(V)) > \Pr(N_1 \geqslant f(V)) = \Pr(N_1 \geqslant T_{12}(n_2 - V)) = p.$$

So, by continuity, one can fix some $c \in (-\varepsilon, \varepsilon)$ such that $p = \Pr(N_1 \geqslant g_c(V))$. Note that

$$\{N_1 \geqslant g_c(V)\} = \{V + N_1 \geqslant c + T_{12}(n_2 - c)\} \text{ and } \{N_1 \geqslant f(V)\} = \{N_1 \geqslant T_{12}(n_2 - V)\},$$

and the former equality depicts a threshold strategy which strictly differs from the latter screening condition $N_1 \geqslant T_{12}(n_2 - V)$. Though both maintain the same capacity of p, the single-crossing property of f and g_c along with the fact that $f' > -1 = g'_c$, suggest that the screening condition $N_1 \geqslant g_c(V)$ omits lower values of V in-exchange to higher ones, relative to the screening condition $N_1 \geqslant f(V)$. Thus, we get

$$\mathbb{E}[V|V + N_1 \geqslant c + T_{12}(n_2 - c)] > \mathbb{E}[V|N_1 \geqslant T_{12}(n_2 - V)] = \mathbb{E}[V|V + N_2 \geqslant n_2],$$

and the first statement of the lemma holds.

To prove the second statement, fix any (V, p). Consider the screening problems $\text{SP}_i = (V, N_i, p)$ and threshold strategies $\hat{\sigma}_{\text{SP}_i}$ for every i. Denote the threshold value of $\hat{\sigma}_{\text{SP}_i}$ by n_i for every i. Thus,

$$\begin{aligned} \mathbb{E}[V|\hat{\sigma}_{\text{SP}_2}(V + N_2) = 1] &= \mathbb{E}[V|V + N_2 \geqslant n_2] \\ &= \mathbb{E}[V|N_2 \geqslant n_2 - V] \\ &= \mathbb{E}[V|T_{12}(N_2) \geqslant T_{12}(n_2 - V)] \\ &= \mathbb{E}[V|N_1 \geqslant T_{12}(n_2 - V)]. \end{aligned}$$

As before, we consider the functions $f(v) = T_{12}(n_2 - v)$ and $g_c(v) = -v + c + T_{12}(n_2 - c)$ defined for every $(v, c) \in \text{Supp}(V)$. Following the same continuity argument (replacing $-\varepsilon$ and ε with sufficiently low and high values, respectively), one can fix c such that $\{N_1 \geqslant g_c(V)\} = \{V + N_1 \geqslant c + T_{12}(n_2 - c)\}$ and both events are of probability p. In other words, c is fixed such that $n_1 = c + T_{12}(n_2 - c)$ and $\{\hat{\sigma}_{\text{SP}_1}(V + N_1) = 1\} = \{N_1 \geqslant g_c(V)\}$. The fact that the single-crossing property still holds and $f' \geqslant g_c'$, ensure again that the threshold strategy $\hat{\sigma}_{\text{SP}_1}$ preforms at least as good as the screening condition $\{N_1 \geqslant f(V)\} = \{\hat{\sigma}_{\text{SP}_2}(V + N_2) = 1\}$. Hence, we conclude that $\widehat{\Pi}_{(V,N_1,p)} \geqslant \widehat{\Pi}_{(V,N_2,p)}$ as needed. ∎

Proof of Theorem 2. We start by showing that S-dominance implies that T_{12} is a contraction. Assume, by contradiction, that T_{12} is not a contraction, so there exists a point n such that $T_{12}'(n) > 1$. Recall that T_{12} is the inverse function of T_{21}^{-1}, so the last inequality suggests that there exists a point m such that $T_{21}'(m) < 1$. By Lemma 4, we deduce that there exists (V, p) such that $\widehat{\Pi}_{(V,N_2,p)} > \widehat{\Pi}_{(V,N_1,p)}$ which contradicts the S-dominance of N_1 over N_2. Thus, we can conclude that T_{12} is indeed a contraction.

Let us now prove the second direction, by assuming that T_{12} is a contraction and establishing the S-dominance of N_1 over N_2. Since N_1 and N_2 are two distinct noise variables (namely, symmetric around zero and independent) and by the fact that T_{12} is a contraction (i.e., a continuously differentiable function), we deduce that there exists a point n such that $T_{12}'(n) < 1$. Thus, by Lemma 4,it follows that $\widehat{\Pi}_{(V,N_2,p)} > \widehat{\Pi}_{(V,N_1,p)}$ for some (V, p). In addition, the weak inequality $\widehat{\Pi}_{(V,N_2,p)} \geqslant \widehat{\Pi}_{(V,N_1,p)}$ holds for every (V, p) by Lemma 4, thus concluding the proof of Theorem 2. ∎

A.3 Proof of Theorem 3

Proof. Fix an impact variable V, a capacity $p \in (0, 1)$, and two noises N_1 and N_2 such that N_2 is a noisy amplification of N_1. Denote $\text{SP}_i = (V, N_i, p)$ for every $i = 1, 2$. We shall prove that $\Pi^*_{\text{SP}_1} \geqslant \Pi^*_{\text{SP}_2}$.

For every noise variable N_i, define the function $f_i(s) = \mathbb{E}[V|V + N_i = s]$. In words, the function f_i produces the expected value of V conditional on a signal s (i.e., on an event $\{V + N_i = s\}$). Since

p is fixed, the optimal strategy $\sigma^*_{\mathrm{SP}_i}$ dictates that $\sigma^*_{\mathrm{SP}_i}(s) = 1$ if $f_i(s) \geqslant t_i$ for some t_i which depends on p and on the distribution of $V + N_i$. Otherwise, if there exist two (positive-probability) sets of signals A and B such that, for every $a \in A$ and $b \in B$, it follows that $\sigma^*_{\mathrm{SP}_i}(a) = 1 > 0 = \sigma^*_{\mathrm{SP}_i}(b)$ and $f_i(a) < f_i(b)$, then $\sigma^*_{\mathrm{SP}_i}$ would not be optimal. Namely, the DM can alternate $\sigma^*_{\mathrm{SP}_i}$ by rejecting signals from A and accepting signals from B (maybe partially, to balance the acceptance ratio) and strictly improve the screening. To exactly sustain the capacity p, the decision maker may need to randomize in case of atoms where $\Pr(V + N_i = s)$ and $\mathbb{E}[V|V + N_i = s] = t_i$. In such cases, the strategy would accept the threshold value with the needed proportion, and otherwise reject the valuations to sustain p.

Define the event $S_i = \{\sigma^*_{\mathrm{SP}_i}(V + N_i) = 1\}$ where $\Pr(S_i) = p$, and denote $q = \Pr(S_1 \cap S_2)$. Observe that $\Pi^*_{\mathrm{SP}_1} = \mathbb{E}[V|S_1] = \frac{q}{p}\mathbb{E}\left[V|S_1 \cap S_2\right] + \frac{1}{p}\mathbb{E}\left[V\mathbf{1}_{S_1\setminus S_2}\right]$. Let us consider the second term, and use the law of iterated expectation (conditional on $V + N_1$) to get

$$
\begin{aligned}
\mathbb{E}\left[V\mathbf{1}_{S_1\cap S_2^c}\right] &= \mathbb{E}[\mathbb{E}[V\mathbf{1}_{S_1}\mathbf{1}_{S_2^c}|V + N_1]] \\
&= \mathbb{E}[\mathbb{E}[V\mathbf{1}_{S_1}|V + N_1]\mathbb{E}[\mathbf{1}_{S_2^c}|V + N_1]] \\
&\geqslant \mathbb{E}[t_1\mathbf{1}_{S_1}\mathbb{E}[\mathbf{1}_{S_2^c}|V + N_1]] \\
&= t_1\mathbb{E}[\mathbb{E}[\mathbf{1}_{S_1}\mathbf{1}_{S_2^c}|V + N_1]] \\
&= t_1\mathbb{E}[\mathbf{1}_{S_1\cap S_2^c}] = t_1(p - q),
\end{aligned}
$$

where we used the fact that, conditional on $V + N_1$, the random variables $V\mathbf{1}_{S_1}$ and $\mathbf{1}_{S_2^c}$ are independent (note that S_2 depends solely on $V + N_1 + N_3$ as $N_2 \sim N_1 + N_3$, and all variables are mutually independent). Thus, $\Pi^*_{\mathrm{SP}_1} \geqslant \frac{q}{p}\mathbb{E}\left[V|S_1 \cap S_2\right] + t_1\frac{p-q}{p}$. Moving on to $\Pi^*_{\mathrm{SP}_2}$, one can follow a similar computation, using the law of iterated expectation, to get the following upper bound

$$
\begin{aligned}
\Pi^*_{\mathrm{SP}_2} &= \mathbb{E}[V|S_2] \\
&= \tfrac{q}{p}\mathbb{E}\left[V|S_2 \cap S_1\right] + \tfrac{1}{p}\mathbb{E}\left[V\mathbf{1}_{S_2\cap S_1^c}\right] \\
&\leqslant \tfrac{q}{p}\mathbb{E}\left[V|S_2 \cap S_1\right] + t_1\tfrac{p-q}{p}.
\end{aligned}
$$

We conclude that $\Pi^*_{\mathrm{SP}_1} \geqslant \Pi^*_{\mathrm{SP}_2}$, as previously stated.

Let us now show that there exists V and p such that the last inequality is strict. Take a normally distributed impact variable $V \sim N(0, 1)$, a capacity $p \in (0, 1)$, and consider the previously used sets $\{S_i\}_{i=1,2}$ and thresholds levels $\{t_i\}_{i=1,2}$, all adjusted for the chosen V and p. Note that for every value $s \in \mathbb{R}$, the conditional distribution of $V|\{V + N_i = s\}$ is non-atomic, and recall that $N_2 \sim N_1 + N_3$.

Henceforth, the proof consists of two stages: first we will show that $\Pr(S_1^c \cap S_2) > 0$, and then we will prove that $\mathbb{E}[V|S_1 \cap S_2^c] > \mathbb{E}[V|S_1^c \cap S_2]$. Let $a_i = \sup\{s : \Pr(S_i|V + N_i < s) = 0\}$ be the maximal value such that every signal below a_i is rejected. Thus, there exists an $\epsilon_0 > 0$ such that for every i and $\epsilon \in (0, \epsilon_0)$ it follows that $\Pr(S_i | V + N_i \in [a_i, a_i + \epsilon)) = 1$. There are two possible cases to consider: either $\Pr(N_3 < a_2 - a_1) > 0$ or $\Pr(N_3 < a_2 - a_1) = 0$.

If $\Pr(N_3 < a_2 - a_1) > 0$, then for a small $\epsilon \in (0, \epsilon_0)$

$$\begin{aligned}
\Pr(S_2^c \cap S_1) &\geqslant \Pr\left(V + N_2 < a_2, V + N_1 \in \left[a_1, a_1 + \tfrac{\varepsilon}{2}\right)\right) \\
&= \Pr\left(N_3 < a_2 - V - N_1, V + N_1 \in \left[a_1, a_1 + \tfrac{\varepsilon}{2}\right)\right) \\
&\geqslant \Pr\left(N_3 < a_2 - a_1 - \tfrac{\epsilon}{2}, V + N_1 \in \left[a_1, a_1 + \tfrac{\varepsilon}{2}\right)\right) \\
&= \Pr\left(N_3 < a_2 - a_1 - \tfrac{\epsilon}{2}\right) \Pr\left(V + N_1 \in \left[a_1, a_1 + \tfrac{\varepsilon}{2}\right)\right) > 0,
\end{aligned}$$

where the last strict inequality follows from the assumption over the distribution of N_3 and ϵ. Therefore, we conclude that $\Pr(S_2^c \cap S_1) > 0$, which implies $\Pr(S_2 \cap S_1^c) > 0$ since $\Pr(S_1) = \Pr(S_2) = p$.

Otherwise, $\Pr(N_3 < a_2 - a_1) = 0 = 1 - \Pr(N_3 \geqslant a_2 - a_1)$ and, by the symmetry of N_3, it follows that $a_2 - a_1 < 0$. Thus, for a sufficiently small $\varepsilon > 0$ we get

$$\begin{aligned}
\Pr(S_2 \cap S_1^c) &\geqslant \Pr\left(V + N_2 \in \left[a_2, a_2 + \tfrac{\varepsilon}{2}\right), V + N_1 < a_1\right) \\
&= \Pr\left(V + N_1 + N_3 \in \left[a_2, a_2 + \tfrac{\varepsilon}{2}\right), V + N_1 < a_1\right) \\
&= \Pr\left(a_2 - N_3 \leqslant V + N_1 < a_2 - N_3 + \tfrac{\varepsilon}{2}, V + N_1 < a_1\right) \\
&\geqslant \Pr\left(a_2 - N_3 \leqslant V + N_1 < a_2 - N_3 + \tfrac{\varepsilon}{2}, N_3 \geqslant 0\right) > 0,
\end{aligned}$$

where the last inequality holds since $V + N_1$ has full support over $\mathbb{R}$ and $\Pr(N_3 \geqslant 0) > 0.5$. Hence, we have shown that $\Pr(S_2 \cap S_1^c) > 0$.

We move on to the second part. Assume that $f_1(s) = \mathbb{E}[V | V + N_1 = s]$ is a non-constant function of the signal $s \in \mathbb{R}$. Then, there exists $p_1 \in (0, 1)$ such that for every capacity $p_0 > p_1$,

$$\mathbb{E}[V | \sigma^*_{(V,N_1,p_1)}(V + N_1) = 1] > \mathbb{E}[V | \sigma^*_{(V,N_1,p_0)}(V + N_1) = 1].$$

This holds by a straightforward convergence-to-the-mean argument, since a more selective and limited choice of values increases the expected value of V relative to an increased capacity, which necessarily introduces sub-optimal valuations. In other words, additional valuations of V are accepted (under capacity p_0 relative to p_1), and the conditional expected value of V subject to these valuations is strictly lower. So, if indeed $f_1(s) = \mathbb{E}[V | V + N_1 = s]$ is a non-constant function, one can fix the capacity p such that $\mathbb{E}[V | S_1 \cap S_2^c] > \mathbb{E}[V | S_1^c \cap S_2]$, as signals outside S_1 yield a strictly lower expected value than the ones in S_1 (and, as was already shown, $\Pr(S_1 \cap S_2^c) = \Pr(S_1^c \cap S_2) > 0$). Therefore, by

Lemma 5 which follows, we conclude that

$$
\begin{aligned}
\Pi^*_{\mathrm{SP}_1} &= \mathbb{E}[V|S_1] \\
&= \tfrac{q}{p}\mathbb{E}\left[V|S_1 \cap S_2\right] + \tfrac{p-q}{p}\mathbb{E}\left[V|S_1 \cap S_2^c\right] \\
&> \tfrac{q}{p}\mathbb{E}\left[V|S_1 \cap S_2\right] + \tfrac{p-q}{p}\mathbb{E}\left[V|S_1^c \cap S_2\right] \\
&= \mathbb{E}[V|S_2] = \Pi^*_{\mathrm{SP}_2},
\end{aligned}
$$

as needed. ■

Lemma 5. *For every impact variable V and noise variable N, the function $f(s) = \mathbb{E}[V|V+N=s]$ is non-constant.*

Proof. Fix an impact variable V and a noise variable N. Assume, with no loss of generality, that $\mathbb{E}[V] = 0$. Note that V is non-degenerate (by definition), so one can fix a small $\epsilon > 0$ such that $\Pr(V > \epsilon)\Pr(V < -\epsilon) > 0$. Take $s \geqslant 0$ such that $\Pr\left(N \in (s-\epsilon, s+\epsilon)\right) > 0$, and denote $I = (s-\epsilon, s+\epsilon)$. Clearly, $\Pr(V + N \geqslant s) \in (0,1)$, and for every $n \in I$, we get $-\varepsilon < s - n < \epsilon$. Thus,

$$\mathbb{E}\left[V|V+n \geqslant s\right] = \mathbb{E}\left[V|V \geqslant s-n\right] > 0 = \mathbb{E}[V].$$

The strict inequality follows from the fact that only low values of V (below $-\epsilon$) are omitted with strictly positive probability. By conditioning on N,

$$\mathbb{E}[V|V+N \geqslant s] = \mathbb{E}\left[\mathbb{E}[V|V+N \geqslant s, N]\right] > 0 = \mathbb{E}[V],$$

and the strict inequality follows from a convex combination of strictly positive and non-negative values. Since $\lim_{s\to-\infty}\mathbb{E}[V|V+N \geqslant s] = \mathbb{E}[V] = 0$, we conclude that $f(s) = \mathbb{E}[V|V+N=s]$ is a non-constant function. ■

A.4 Proof of Lemma 1

Proof. Without loss of generality, assume that $N \sim U[0,1]$ and denote $\mathrm{Supp}(V) = [\underline{V}, \overline{V}]$. Fix two signals $s_1 > s_2$ where $s_i \in \mathrm{Supp}(V+N)$ for every i. We will show that $\mathbb{E}[V|V+N=s_1] \geqslant \mathbb{E}[V|V+N=s_2]$. If that is the case, then for any two sets A and B such that $\Pr(V+N \in A)\Pr(V+N \in B) > 0$ and A is point-wise strictly above B, we maintain the same monotonic relation of $\mathbb{E}[V|V+N \in A] > \mathbb{E}[V|V+N \in B]$ and the statement follows.

Note that N is uniformly distributed on $[0,1]$, so the random variable $V+N$ has a non-atomic distribution and

$$\mathrm{Supp}(V|\{V+N=s_i\}) = \left[\max\{s_i - 1, \underline{V}\}, \min\{s_i, \overline{V}\}\right].$$

Since N supports all points in $[0,1]$ with equal weight, one can verify that the projection of $V+N=s_i$ onto V maintains the distribution of V, conditional on the same support, such that

$$V|\{V+N=s_i\} \sim V|\{V \in \left[\max\{s_i-1,\underline{V}\},\min\{s_i,\overline{V}\}\right]\}.$$

Therefore, the deviation from s_2 to s_1 increases (potentially weakly) the bounds $\max\{s_i-1,\underline{V}\}$ and $\min\{s_i,\overline{V}\}$, which ensures that the inequality $\mathbb{E}[V|V+N=s_1] \geqslant \mathbb{E}[V|V+N=s_2]$ holds. ■

A.5 Proof of Lemma 2

Proof. Consider the screening problems $\mathrm{SP}_i=(V,N_i,p)$ for every i. Denote the optimal threshold strategy $\hat{\sigma}_{\mathrm{SP}_i}$, and let s_i be the threshold value such that $\hat{\sigma}_{\mathrm{SP}_i}(s)=\mathbf{1}_{\{s\geqslant s_i\}}$. Denote the events $A_i=\{V+N_i\geqslant s_i\}$, and probabilities $p=\Pr(A_1)=\Pr(A_2)$, $p'=\Pr(A_1\cap A_2)$.

We begin by showing that $\mathbb{E}[V|A_1\cap A_2^c]\geqslant \mathbb{E}[V|A_1^c\cap A_2]$. The lines $V+N_1=s_1$ and $V+\lambda N_1=s_2$ intersect at $(V,N_1)=\left(t_1-\frac{s_1-s_2}{1-\lambda},\frac{s_1-s_2}{1-\lambda}\right)$, and

$$A_1\cap A_2^c=\left\{V>s_1-\tfrac{s_1-s_2}{1-\lambda}, N_1\in\left[s_1-V,\tfrac{s_2-V}{\lambda}\right)\right\}\ ,\ \ A_1^c\cap A_2=\left\{V<s_1-\tfrac{s_1-s_2}{1-\lambda}, N_1\in\left[\tfrac{s_2-V}{\lambda},s_1-V\right)\right\}.$$

So, in terms of V, we get a point-wise dominance given $A_1\cap A_2^c$ relative to $A_1^c\cap A_2$, and $\mathbb{E}[V|A_1\cap A_2^c]\geqslant \mathbb{E}[V|A_1^c\cap A_2]$. Therefore,

$$\begin{aligned}
\mathbb{E}[V|\sigma^*_{s_2}(V+N_2)=1] &= \mathbb{E}[V|A_2]\\
&= \tfrac{p'}{p}\mathbb{E}\left[V|A_1\cap A_2\right]+\tfrac{p-p'}{p}\mathbb{E}\left[V|A_1^c\cap A_2\right]\\
&\leqslant \tfrac{p'}{p}\mathbb{E}\left[V|A_1\cap A_2\right]+\tfrac{p-p'}{p}\mathbb{E}\left[V|A_1\cap A_2^c\right]\\
&= \mathbb{E}[V|A_1]=\mathbb{E}[V|\sigma^*_{s_1}(V+N_1)=1].
\end{aligned}$$

Note that the inequality becomes strict whenever the two threshold strategies do not trivially coincide ($p>p'$), and the statement holds. ■

A.6 Proof of Lemma 3

Proof. Assume, by contradiction, that there exists a random variable N, independent of N_1, such that $N_1+N\sim N_2\sim U[0,1.5]$. Evidently, $\mathrm{Supp}(N)\subseteq[0,0.5]$, otherwise $\mathrm{Supp}(N_1+N)\neq[0,1.5]$ as needed. By conditioning on N_1, we get

$$\frac{1}{3}=F_{N_1+N}\left(\tfrac{1}{2}\right)=\int_0^1 \Pr\left(N\leqslant \tfrac{1}{2}-n\right)dn=\int_{-\frac{1}{2}}^{\frac{1}{2}}\Pr\left(N\leqslant k\right)dk=\int_0^{\frac{1}{2}}\Pr\left(N\leqslant k\right)dk,$$

and

$$\begin{aligned}
F_{N_1+N}(1) &= \int_0^1 \Pr(N \leqslant 1-n)\, dn = \int_0^1 \Pr(N \leqslant k)\, dk = \int_0^{\frac{1}{2}} \Pr(N \leqslant k)\, dk + \int_{\frac{1}{2}}^1 1 dk \\
&= F_{N_1+N}\left(\tfrac{1}{2}\right) + \frac{1}{2} = \frac{1}{3} + \frac{1}{2} = \frac{5}{6},
\end{aligned}$$

contradicting the preliminary assumption which suggests that $F_{N_1+N}(1) = F_{N_2}(1) = \frac{2}{3}$. ■