# Screening Dominance: A Comparison of Noisy Signals*

David Lagziel† and Ehud Lehrer‡

October 27, 2020

Abstract:

This paper studies the impact of noisy signals on screening processes. It deals with a decision problem in which a 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". We compare different noisy signals under threshold strategies and optimal ones, and provide several characterizations of cases in which one noise is preferable over another. Accordingly so, we establish a novel method to compare noise variables using a contraction mapping between percentiles.

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

Keywords: screening dominance; 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 the Adam Smith Business School Microtheory Seminar, the Manchester Economic Theory Seminar, ORSIS 2019, the Tel Aviv University Game Theory and Mathematical Economics Research Seminar, the Technion Game Theory Seminar, and the Bar Ilan University Game Theory 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 have just been appointed to a top managing position in your field, say, the editor of a top-tier academic journal. In an effort to promote your journal, you turn to the editorial board for suggestions concerning ways to improve the journal's screening process. Now an editorial member approaches you claiming that there is a rather simple method to strictly improve the screening. Specifically, the member advises you to use a "lucky coin toss" whose result would be weighed into the evaluation process. Hearing this, you will quite possibility consider the suggestion to be a joke, and for good reasons. It does seem absurd that one could improve a screening by introducing additional independent noise to the process. Nonetheless, in this paper we substantiate the potential superiority of the mentioned proposal, and we do so by studying the impact of different noisy signals 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 DM could be, for example, 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, obviously, 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 a basic screening model, 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 $V + N$. The DM strives 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 a few obvious reasons: threshold strategies are simple, commonly used, and, in the absence of noise, they are indeed optimal. In this paper, we also adhere to this line of thinking. Indeed, our first observation 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 of generating "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, and 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.

This preliminary observation is merely the overture to a much broader question concerning the way different noisy signals 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, while holding the capacity fixed.

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

The next stage of our analysis focuses on optimal screening strategies. Assuming that optimal screening strategies are applied, we prove that 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 and independent noise. This characterization strongly relates to our contraction notion, since the mentioned condition (for normal distributions) is equivalent to $N_1$ being a contraction of $N_2$.

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. In this class of noises, we show that screening dominance is not fully characterized by additive noise, but follows from the previously mentioned contraction property. Overall, our contraction characterization accounts for *optimal screening* under at least two (presumably, the more important) classes of noises: normal and uniform distributions.

## 1.1 Related literature and main contribution

The economic research on screening and noisy signals ranges from job-market signaling and education to insurance and credit markets (see, e.g., Spence (1973), Stiglitz (1975), Rothschild and Stiglitz (1976), Stiglitz and Weiss (1981), Sah and Stiglitz (1986, 1988, 1991), and Meyer (1991) among many others). Note that these papers typically focus on costly screening and strategic signaling, while we consider a non-strategic and costless signaling model. So, in the relevant literature, the papers that are closest to ours are those of Blackwell (1951, 1953), Lehmann (1988), Quah and Strulovici (2009), and Rothschild and Stiglitz (1970, 1971).

Starting with the former, Blackwell (1951, 1953) compares different information structures for the purpose of maximizing the expected payoff in a decision problem. In terms of the main focus and classification in the literature, our study is quite close to Blackwell's work. However, there are two key differences between the studies. First, in Blackwell's model, the DM observes noisy signals and aims to maximize his expected utility. In contrast, in our model, the DM strives to maximize the *conditional expectation* of $V$. Second, there is a significant difference concerning the comparison in question. Blackwell compares two information structure defined on the same state-space. We, however, allow the underlying state-space (namely, the impact variable) to vary. In our framework, every noise variable generates different information structures, based on the different impact variables. Thus, we actually compare different *information-structure generators* (the noises) rather than information structures per se. These differences eventually lead to two well-distinct characterizations — Blackwell's garbling notion versus our contraction mapping.

Despite the clear differences in motivation and modeling choices, our main result in this paper, namely the characterization of screening dominance through a contraction mapping, is related to the studies of Lehmann (1988) and Quah and Strulovici (2009) — both in the vast literature of statistical decision theory and information economics. In the earlier study, Lehmann (1988) defines a notion of comparison between two experiments such that one is *more informative* than the other (or *more accurate*, following the terminology of Persico (2000)). In our model of additive noise, one can show that our contraction property and Lehmann's notion of informativeness coincide (see Theorem 5.2 in Lehmann (1988)). Yet, a crucial distinction is that Lehmann requires either strongly unimodal densities (in Theorem 5.2), or a monotone likelihood ratio property (in Theorem 5.1), which are irrelevant in our model. In addition, we provide a characterization in the context of screening problems where the decision maker tries to maximize some goal function, whereas Lehmann adopts a more statistical approach of comparing information structure in general.

A second study in this field, which follows the work of Lehmann (1988) and also relates to ours, is by Quah and Strulovici (2009). In Lemma 3 therein, Quah and Strulovici show that for any threshold strategy in a less informative screening problem (in the sense of Lehmann), one can devise a superior threshold strategy in the more information screening problem. For that purpose they consider a general utility function and a different signaling system (specifically, they do not confine themselves to additive noise, but require a fully supported set of signals in every state of the world). Other than the underlying goal and motivation, the two crucial differences between our result and the results of Quah and Strulovici (2009) are: (i) we provide a characterization (rather than a necessary condition); and (ii) we follow a (binding) capacity constraint on the set of accepted elements, which typically does not appear in this line of research. Evidently, a capacity constraint is rather natural in screening problems, in comparison to general decision problems.

A more recent study in this field is by Di Tillio et al. (2020). They follow the work of Lehmann (1988) and Quah and Strulovici (2009), and ask whether it is beneficial for the DM to observe a given set of i.i.d. signals, or to first strategically select them from a larger set of signals. Similarly to our analysis, their core results focus on threshold strategies and additive noise (called *location experiments*; see Section 4 therein).

Moving on to the work of Rothschild and Stiglitz (1970, 1971), the first aspect that associates our work to theirs is the underlying goal: relating probabilistic properties of random variables to the preferences of a rational decision maker. Rothschild and Stiglitz (1970, 1971) achieve this goal by introducing 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, we consider an additive independent noise and provide several equivalence results between the induced partial order (over noises) and screening dominance.

This first similarity naturally leads to the second important connection between the indicated 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 introduce a contraction mapping 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 similarity between the studies in question 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. To compare, we follow the model of Lagziel and Lehrer (2019) with its broad set of applications, which range from peer-reviewed academic publishing to credit ratings.

### 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 analyzing 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 individual values of the elements 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 acceptance of a specific valuation and 0 denotes rejection.[2] To avoid trivial solutions, we fix an acceptance rate, a *capacity* level $p \in (0, 1)$, which defines the share of accepted elements so that every screening strategy $\sigma$ must ensure that $\Pr(\sigma(V + N) = 1) = p$.

To motivate this model, one can think of the DM as an editor of a peer-reviewed academic journal who approaches referees to evaluate a set of academic papers: $V$ denotes the papers' potential impact, $V + N$ are the referees' evaluations, and $\sigma$ is the editor's decision rule to accept or reject a paper. Other possible 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 the accepted elements, subject to some capacity constraint.

We generally refer to the triplet $\text{SP} = (V, N, p)$ as a *screening problem.* Given a screening

[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 measurability requirements hold.

problem SP and a screening strategy $\sigma$, the expected value of accepted elements is

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

The DM's goal is to maximize $\Pi_{\text{SP}}$. We denote the DM's optimal screening strategy and optimal expected payoff by $\sigma^*_{\text{SP}}$ and $\Pi^*_{\text{SP}}$, 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 analyzing the class of threshold strategies, and this for two main reasons. The first reason 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 $\sigma$ 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 $\hat{\sigma}_{\text{SP}}$ and $\widehat{\Pi}_{\text{SP}}$, 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 $p$, the DM may need to impose a partially random screening such that valuations that are subject to an atom are randomly split. In such cases, one should consider a more general screening strategy where $\sigma : \mathbb{R} \to [0,1]$. 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.

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

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

*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 $N_!$ is at least as good as an optimal screening under $N_2$ (and, in some cases, strictly better), independently of 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 use a weaker notion where 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 pair $(V,p)$, and the inequality is strict for some pair $(V,p)$.

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 a noise variable $N_1$ if $N_2 \sim N_1 + N_3$ and the 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 dominance of one noise over another. We conclude by noting that a noisy amplification is also an MPS (as noise variables have zero mean and the added noise $N_3$ is independent of $N_1$), whereas the converse is not true.

# 3 Screening under threshold strategies

This section is divided into two parts, each presenting one key result of the paper. The first part, in Section 3.1, shows how additional noise can strictly improve a screening process. The second part, in Section 3.2, presents the first 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 Adding noise to a screening problem

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 the result into her decision. On the other hand, how can a simple lottery improve a screening if we are merely introducing random noise into the process? In this section, we shall attempt to resolve this puzzle.

We begin with a straightforward result stating that lucky coins exist, and later motivate it with a simple example. Proposition 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 $N_1$.

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

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

To exemplify this result, consider the following (rather stylized) example. Assume that the SAT scores of a large group of students is distributed uniformly on $[800, 1200]$. In addition, assume that only half of the students are accepted to undergraduate studies and acceptance is set according to some cut-off value. Now, due to some computational error, grades are randomly distorted by $\pm 200$ points, with equal probabilities, so the observed grades are uniformly distributed on $[600, 1400]$. With and without this error, the screening is set from 1000 and above, thus capturing half of the population. However, the error completely distorts the original distribution and the average score of accepted students is 1000, compared to 1100 without the error. In terms of the average score, note that this "noisy" screening is rather useless since it preforms *as good as random acceptance*. However, if the same computational error occurs twice (independently), the aggregate error is distributed according to

$$N = \begin{cases} \pm 400, & \text{with probability } 0.25, \\ 0, & \text{with probability } 0.5. \end{cases}$$

In words, the double error produces an accurate evaluation with probability 0.5, and combined with the original uniform distribution on $[800, 1200]$, the average score of accepted students is 1050. Therefore, screening under these two independent errors is better than screening under a single error (although the former is considered riskier in the sense of Rothschild and Stiglitz).

We will now explain the driving force behind this result. First, recall the key observation of Lagziel and Lehrer (2019) that $\widehat{\Pi}_{(V,N,p)}$ is not necessarily a monotone 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. $p$ is due to 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 significant 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 an unbiased noise can distort the distribution, so that the noisy valuations do not reflect the "true" ordering of the elements' impact. The additional noise can partially rectify this distortion, at least to some extent, by "re-ordering" the values of $V$.

An important component of this result is that we are applying threshold strategies. Should the DM have the ability to apply optimal strategies, 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. In the meantime, let us point out that the following results (specifically, Theorem 1 and Theorem 2) do not depend on the optimality of threshold strategies.

#### 3.1.1 Robustness of Proposition 1

We wish to discuss two robustness concerns regarding the result of Proposition 1.

First, the noises used in the proof of Proposition 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 dependence on $p$ hinges on the need to sustain symmetric noises. Therefore, if one allows for asymmetric noises, the result of Proposition 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 $V$. 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 Proposition 1 remains valid.

Third, though this require some more work, one can also extend Proposition 1 to unbounded impact variables. Specifically, using Lemma 2 of Lagziel and Lehrer (2019), one can construct a screening bias for an unbounded impact variable. Then, to partially correct this bias, one can add an i.i.d. noise, and follow a similar computation as the one given in the previously discussed SAT example.[3]

[3]Note that the SAT example itself could be extended to general distributions (rather than uniform) on a broader support than the current one; this should provide additional intuition for the extension of Proposition 1 to unbounded variables.

### 3.2 A partial characterization of screening dominance

To characterize screening dominance under threshold strategies, we begin by defining a mapping that transforms any non-atomic noise into another non-atomic noise using percentiles translation. This mapping will be used to compare a threshold strategy under one noise variable with 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 our mapping is a contraction or not. If the mapping 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) mapping $T_{ij}$ by

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

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

Let us now review the key properties of the PT mapping. Since both noises are non-atomic with convex supports, the CDFs are strictly increasing on these sets and the mapping $T_{ij}(n)$ is well defined and strictly increasing. Second, it is straightforward to verify that $T_{12}$ is the inverse of $T_{21}$ and both are bijective maps (one-to-one correspondences) between the relevant supports. In that case, $T_{21} = T_{12}^{-1}$ and $T'_{21}(n) = \frac{1}{T'_{12}(T_{21}(n))}$.

To simplify the exposition, we introduce two additional definitions: (i) $N_i$ and $N_j$ are called *continuous* if both noises are non-atomic with convex supports, and if $T_{ij}$ 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 $n \in \text{Supp}(N_j)$ and $T_{ij}$ is continuously differentiable. Roughly speaking, one noise variable is a contraction of another if the percentiles of the former are closer together than the percentiles of the latter.

A comparison between the contraction property and the notion of a MPS is clearly needed. Given our symmetric-noise assumption, one can easily verify that a contracting PT mapping $T_{ij}$ implies that $N_j$ is a MPS of $N_i$,[4] whereas the converse is not true. Note, however, that the mentioned symmetry assumption is not a necessary one for our main results. If this assumption is omitted, then the notion of a MPS does not follow from the contraction property since the PT

[4]Given the symmetry assumption, MPS is also equivalent to second-order stochastic dominance.

mapping can transform a symmetric noise variable into an asymmetric one (with a zero mean) such that the original noise variable is not a MPS of the transformed one. We thus conclude that the two notions are basically distinct and complementary. On the other hand, and as already discussed in Subsection 1.1, our contraction property does coincide with the *more-informative* notion of Lehmann (1988). We explicitly establish this connection in Appendix A.2.

As follows from Theorem 1 below, the contraction property of the TP mapping 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 1.** *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.*

Theorem 1 joins a long list of results that relate decision making to probabilistic properties, which are later applied to various economic models (see, e.g., Holmstrom (1979) and Grossman and Hart (1983), among others). Many of these studies trace back to the work of Blackwell (1951, 1953), who characterized an information superiority through the garbling of signals.

The main focus of our study is the information needed to maximize the conditional expectation of $V$. In Blackwell's model, in contrast, the decision maker's objective is to maximize the (unconditional – over the entire state-space) expected state-dependent utility. The application of different screening strategies to different screening problems typically induces different acceptance sets, which do not allow mean-spread comparison (garbling) as required by Blackwell's characterization.

Another significant difference between Blackwell's work and ours is that Blackwell deals with information structures that associate a noisy signal to every state. He then compares two such structures, both defined on the *same* state-space. Here, the underlying state-space (in our context, the impact variables $V$) varies. Indeed, given $V$, a noise $N$ generates an information structure defined over $V$. However, the same noise $N$ may generate an information structure over every other impact variable. Allowing all possible impact variables, we actually compare different noises as *information-structure generators*, rather than information structures.[5]

Due to these two significant differences, it is only natural that our characterization through the notion of noise contraction is not directly connected to Blackwell's garbling. In particular

[5]Related to this goal, one can find some similarity in a recent study by Di Tillio et al. (2020) in which the DM needs to decide between a random sample of i.i.d. signals, or a strategically chosen sub-sample of i.i.d. signals.

our characterization cannot be derived from that of Blackwell, and vice versa.

Figure 1 provides intuition for the proof of Theorem 1. The graph on the left represents threshold-screening under $N_2$, where $l_2$ denotes the threshold line for that screening. Using the PT mapping, one can translate $l_2$ to terms of $N_1$, thus obtaining the $l_2'$ line in the right 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 with respect to $N_1$ (given by the grey areas in the right figure and the $l_1$ 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.

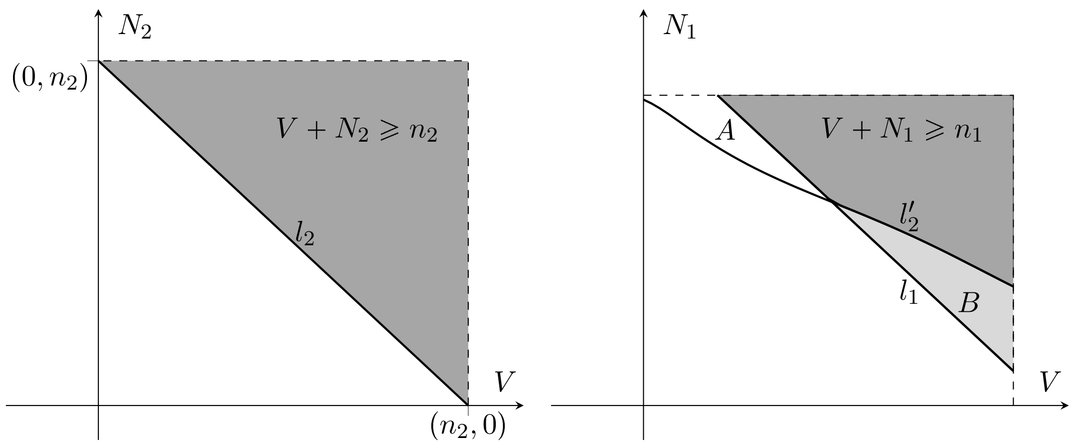


Figure 1: Each graph $l_i : V + N_i = n_i$ represents threshold-screening in the screening problem $\text{SP}_i$, and each shaded area $V + N_i \geqslant n_i$ represents the accepted valuations. Line $l_2' : N_1 = T_{12}(n_2 - V)$ is the representation of $l_2$ in terms of $N_1$ using the PT mapping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

What happens if the PT mapping is not a contraction? Assuming that noises are distinct and unbiased, then $T_{12}$ is not a linear mapping. Thus, there exists a point $n$ such that $T_{12}'(n) > 1$, which suggests that $T_{12}$ is locally expanding in the neighborhood of $n$ (recall that the PT mapping 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 1, but when translating $N_1$ to $N_2$. This guarantees that there exist an impact variable and a 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 mapping is provided by 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.

**Remark 1.** *Note that the statement of* Theorem 1 *holds whether threshold strategies are optimal or not. This stands in contrast to* Proposition 1*, where threshold strategies are suboptimal. Moreover, in* Section 4.1 *we extend* Theorem 1 *to* S-*dominance* (*under optimal strategies*) *given normally distributed noises* (*see* Observation 1 *below*).

### 3.2.1 Contraction may not imply S-dominance (under optimal strategies)

We conclude this section with an explanation concerning the problems that arise when reverting to optimal (potentially, non-threshold) strategies. 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 the noise $N_2$, which would generate four distinct signals $s \in \{-2, 0, 2, 4\}$.

In order to compare this example to Theorem 1, take a capacity of $p = 0.5$ and note that

$$\Pi^*_{\mathrm{SP}_2} = 2 > 1.5 =\geqslant \Pi^*_{\mathrm{SP}_1}.$$

That is, an optimal screening in $\mathrm{SP}_2 = (V, N_2, p)$ produces a strictly higher expected value than an optimal screening under $\mathrm{SP}_1 = (V, N_1, p)$, although $N_1$ is a contraction of $N_2$. In other words, the information ambiguity that $N_1$ produces in $\mathrm{SP}_1$ implies that $N_1$ does not S-dominate $N_2$, and it also explains the need to take a different approach once dealing with optimal strategies, as we do in the following section.

## 4 Screening under optimal strategies

In this section we focus on screening dominance under optimal strategies, for which purpose we divide our analysis into two parts. In Section 4.1, we show that a noisy amplification is a sufficient condition for screening dominance. Moreover, if we restrict attention to normally distributed noises, we prove that the noisy amplification condition is, in fact, a characterization of screening dominance. Then, 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 such noises, whereas the contraction result of Theorem 1 does provide a necessary condition under uniform 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 2 below states that a noisy amplification of one noise variable is dominated, in terms of screening, by that variable.

**Theorem 2.** *If $N_2$ is a noisy amplification of $N_1$, then $N_1$* S-*dominates $N_2$.*

In order to prove Theorem 2, one needs to devise an optimal screening strategy for general screening problems. The proof of Theorem 2 does that based on the function

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

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

**Remark 2.** *Note that* S-dominance *is based on the* strict *superiority of $N_1$ over the noisy amplification $N_2$ for some impact variable $V$ and capacity $p$. Thus, the construction and implementation of an optimal screening strategy is necessary for the proof of* Theorem 2. *Specifically, given $\mathrm{SP}_1 = (V, N_1, p)$, one cannot simply assume that the* DM *replicates the optimal screening strategy as in $\mathrm{SP}_2 = (V, N_2, p)$, by using the original signal $V + N_1$ and a proper randomization rule. The latter merely produces the same expected value as in $\mathrm{SP}_2$.*

An immediate conclusion from Theorem 2 is the 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 2, so we need only consider the other direction, starting with the screening dominance of $N_1$ over $N_2$. If $N_1$ S-dominates $N_2$, there exist 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 2) that $N_2$ is a noisy amplification of $N_1$.

Corollary 1 along with the basic properties of normal distributions allow us to derive the following observation regarding noise contraction under optimal strategies. Specifically, for normally distributed noise variables, $N_2$ is a noisy amplification of $N_1$ if and only if $N_1$ is a contraction of $N_2$. Hence, the characterization given in Corollary 1 applies for the contraction property, as well.

**Observation 1.** *Fix two normally distributed noise variables $N_1$ and $N_2$. Then, $N_1$ is a contraction of $N_2$ if and only if $N_1$ S-dominates $N_2$.*

This observation extends Theorem 1 and the counterexample given in Section 3.2.1, by showing that, for some classes of noises, the contraction property does lead to S-dominance (under optimal strategies). We pursue this goal in the following section.

### 4.2 Screening dominance under uniform noises

Corollary 1 naturally raises the following 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 noises with convex support.

We begin our analysis by establishing, in Lemma 1 below, that under uniformly distributed noises (with convex supports) threshold strategies are optimal.

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

In view of 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 analyze 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 $N_1$ S-dominates $N_2$ under (the optimal) threshold strategies. In the following lemma, we prove this result for general distributions without confining ourselves to continuous noises.

**Lemma 2.** *Fix two screening problems $\mathrm{SP}_i = (V, N_i, p)$, $i = 1, 2$, such that $N_2 \sim cN_1$ 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 of 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[6] noises $N_1 \sim U[0,1]$ and $N_2 \sim U[0,3/2]$. Lemma 1 states that the optimal screening strategy in any screening problem (under these noises) is a threshold strategy. Lemma 2 establishes that $N_1$ dominates $N_2$ since $N_2 \sim (3/2)N_1$. Lemma 3 below 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 the other, while the noisy-amplification condition is violated.

**Lemma 3.** *If $N_1 \sim U[0,1]$ and $N_2 \sim U[0,3/2]$, then $N_2$ is not a noisy amplification of $N_1$.*

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

# 5 Conclusion

In this paper we provided several novel insights into 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. 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 of screening dominance among different types of noises, and most importantly, our main characterization result shows that some form of contraction among the noises' distributions is essential for screening dominance.

# References


Blackwell, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, Calif.: University of California Press, 93–102.

——— (1953): "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 24, 265–272.

Di Tillio, A., M. Ottaviani, and P. Norman Sorensen (2020): "Strategic Sample Selection," .


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

Grossman, S. J. and O. D. Hart (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, 51, 7.

Holmstrom, B. (1979): "Moral Hazard and Observability," *The Bell Journal of Economics*, 10, 74.

Lagziel, D. and E. Lehrer (2019): "A Bias of Screening," *American Economic Review: Insights*, 1, 343—-356.

Lehmann, E. L. (1988): "Comparing Location Experiments," *The Annals of Statistics*, 16, 521–533.

Meyer, M. A. (1991): "Learning from Coarse Information: Biased Contests and Career Profiles," *The Review of Economic Studies*, 58, 15.

Persico, N. (2000): "Information Acquisition in Auctions," *Econometrica*, 68, 135–148.

Quah, J. K.-H. and B. Strulovici (2009): "Comparative Statics, Informativeness, and the Interval Dominance Order," *Econometrica*, 77, 1949–1992.

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, 629–649.

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.

Sah, R. K. and J. E. Stiglitz (1986): "The Architecture of Economic Systems: Hierarchies and Polyarchies," .

——— (1988): "Economics of Committees," *Economic Journal*, 98, 451–470.

——— (1991): "The Quality of Managers in Centralized Versus Decentralized Organizations," *The Quarterly Journal of Economics*, 106, 289–295.

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 Proposition 1

**Proof.** Fix an impact variable $V$ and a capacity $p \in (0, 1)$. Assume, without loss of generality, that $V$ is supported on $[0, 1]$. We examine separately three cases: $p < 0.5$, $p = 0.5$, and $p > 0.5$. In general, denote the screening problem by $\mathrm{SP}_i = (V, N_i, p)$ for every $i$ and every noise $N_i$.

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

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

Clearly, $\mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_1}(V + N_1) = 1] = \mathbb{E}[V]$. 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 when $p > 0.5$, perform a similar computation with $p$ replaced by $1-p$. 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 when $p = 0.5$, set $N_1 = \pm 0.6$ with equal probabilities, and set $N_3 = \pm 0.2$ 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 The contracting PT mapping and Lehmann's ordering of signals

Lehmann (1988) defines an ordering of signals in the following manner. Consider two noisy signals $X_1$ and $X_2$ about some impact variable $V$. Note that we consider additive noises, so Lehmann's signals translate to $X_i = V + N_i$, for every $i$, under our terminology. Let $G_i(\cdot|v)$ be the conditional distribution of $X_i$, given $V = v$, and consider the mapping $h_v(x) = G_1^{-1}(G_2(x|v)|v)$ by Lehmann (1988); see Theorem 5.1 therein. Lehmann assumes that these conditional distributions are differentiable and maintain the monotone likelihood ratio property (MLRP) for every $v$, to establish that $X_1$ is *more informative*[7] than $X_2$ if and only if $h_v(x)$ is non-decreasing in $v$, for every $x$.

To explicitly relate this informativeness notion to our contracting PT mapping, note that $G_i(x|v) = F_i(x - v)$ under an additive-noise set-up, where $F_i$ is the CDF of the noise variable $N_i$. Thus, $G_1^{-1}(y|v) = F_1^{-1}(y) + v$ and $h_v(x)$ translates to

$$h_v(x) = G_1^{-1}(G_2(x|v)|v) = F_1^{-1}(G_2(x|v)) + v = F_1^{-1}(F_2(x - v)) + v = T_{12}(x - v) + v.$$

Therefore, under all needed differentiability assumptions, $h_v(x)$ is non-decreasing in $v$ for every $x$ if and only if $T'_{12}(n) \leqslant 1$ for every $n$.

Lehmann also considers an additive-noise set-up and assumes that the densities are strongly unimodal (see Theorem 5.2 therein), to prove that one signal is more informative than another if the updated $h_v(\cdot)$ mapping is contracting. Therefore, the connection is straightforward. Yet, one should note that the characterizations remain distinct as we require neither the MLRP, nor strongly unimodal densities, whereas we do require a continuously differentiable PT mapping.

## A.3 Proof of Theorem 1

For the proof of Theorem 1 we need 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.** 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 $I = (n_2 - \varepsilon, n_2 + \varepsilon)$ such that $T'_{12}(n) < 1$ 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 have

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

[7] Lehmann uses the term *more effective* when restricting the discussion to a subset of decision problems.

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 because $T_{12}$ is strictly increasing.

Consider the function $f(v) = T_{12}(n_2 - v)$ for $v \in (-\varepsilon, \varepsilon)$. Clearly, $f$ is strictly decreasing, differentiable, and $f'(v) = -T'_{12}(n_2 - v) > -1$ 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 graphs of the functions $f(v)$ and $g_c(v)$ intersect exactly once, at $(c, T_{12}(n_2 - c))$. Specifically, $g_\varepsilon(v) \geqslant f(v)$, while $g_{-\varepsilon}(v) \leqslant f(v)$.

We can now use $g_c$ to construct a threshold (screening) strategy for the 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 $p$, the single-crossing property of $f$ and $g_c$ together 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 for 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 $\mathrm{SP}_i = (V, N_i, p)$ and threshold strategies $\hat{\sigma}_{\mathrm{SP}_i}$ for every $i$. Denote the threshold value of $\hat{\sigma}_{\mathrm{SP}_i}$ by $n_i$

for every $i$. Thus,

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

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 $(v, c) \in \mathrm{Supp}(V)$. Following the same continuity argument (replacing $-\varepsilon$ and $\varepsilon$ 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 so $n_1 = c + T_{12}(n_2 - c)$ and $\{\hat{\sigma}_{\mathrm{SP}_1}(V + N_1) = 1\} = \{N_1 \geqslant g_c(V)\}$. The fact that the single-crossing property still holds and the inequality $f' \geqslant g_c'$ ensure again that the threshold strategy $\hat{\sigma}_{\mathrm{SP}_1}$ performs at least as well as the screening condition $\{N_1 \geqslant f(V)\} = \{\hat{\sigma}_{\mathrm{SP}_2}(V + N_2) = 1\}$. Hence, we conclude that $\widehat{\Pi}_{(V,N_1,p)} \geqslant \widehat{\Pi}_{(V,N_2,p)}$, as needed. ∎

**Proof of Theorem 1.** 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 mapping of $T_{21}^{-1}$, so the last inequality suggests that there exists a point $m$ such that $T_{21}'(m) < 1$. By Lemma 4, there exists a pair $(V, p)$ 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: assuming that $T_{12}$ is a contraction, we establish 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 since $T_{12}$ is a contraction (and so a continuously differentiable mapping), we deduce that there exists a point $n$ such that $T_{12}'(n) < 1$. Thus, by Lemma 4, we see 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. ∎

## A.4 Proof of Theorem 2

**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 $\mathrm{SP}_i = (V, N_i, p)$ for $i = 1, 2$. We shall prove that $\Pi^*_{\mathrm{SP}_1} \geqslant \Pi^*_{\mathrm{SP}_2}$.

For the noise variable $N_i$, define the function $f_i(s) = \mathbb{E}[V|V + N_i = s]$, $i = 1, 2$. 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 $\sigma^*_{\mathrm{SP}_i}(a) = 1 > 0 = \sigma^*_{\mathrm{SP}_i}(b)$ and $f_i(a) < f_i(b)$ for every $a \in A$ and $b \in B$, 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 DM may need to randomize in case atoms are present 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 carry out a similar computation, using the law of iterated expectation, to get the following upper bound:

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

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

Now, let us show that there exist $V$ 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 $N_2 \sim N_1 + N_3$.

Next, the proof consists of two stages: first we will show that $\Pr(S_1^{\mathrm{c}} \cap S_2) > 0$, and then that $\mathbb{E}[V|S_1 \cap S_2^{\mathrm{c}}] > \mathbb{E}[V|S_1^{\mathrm{c}} \cap S_2]$. 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 $\epsilon \in (0, \epsilon_0)$ one has $\Pr(S_i|V + N_i \in [a_i, a_i + \epsilon)) = 1$. There are two possible cases to consider: $\Pr(N_3 < a_2 - a_1) > 0$, and $\Pr(N_3 < a_2 - a_1) = 0$.

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

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

where the last strict inequality follows from the assumptions on the distributions of $N_3$ and $\epsilon$. Therefore, $\Pr(S_2^{\mathrm{c}} \cap S_1) > 0$, which implies $\Pr(S_2 \cap S_1^{\mathrm{c}}) > 0$, 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

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

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 $\Pr(S_2 \cap S_1^{\mathrm{c}}) > 0$.

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 $\mathbb{E}[V|S_1 \cap S_2^{\mathrm{c}}] > \mathbb{E}[V|S_1^{\mathrm{c}} \cap S_2]$, as signals outside $S_1$ yield a strictly lower expected value than the ones in $S_1$ (and, as already shown, $\Pr(S_1 \cap S_2^c) = \Pr(S_1^c \cap S_2) > 0$). Therefore, by Lemma 5 below, we conclude that

$$\begin{aligned}
\Pi^*_{SP_1} &= \mathbb{E}[V|S_1] \\
&= \frac{q}{p}\mathbb{E}\left[V|S_1 \cap S_2\right] + \frac{p-q}{p}\mathbb{E}\left[V|S_1 \cap S_2^c\right] \\
&> \frac{q}{p}\mathbb{E}\left[V|S_1 \cap S_2\right] + \frac{p-q}{p}\mathbb{E}\left[V|S_1^c \cap S_2\right] \\
&= \mathbb{E}[V|S_2] = \Pi^*_{SP_2},
\end{aligned}$$

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 $\Pr(N \in (s - \epsilon, s + \epsilon)) > 0$, 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.5 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 monotone relation $\mathbb{E}[V|V + N \in A] > \mathbb{E}[V|V + N \in B]$, 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 $V$ preserves the distribution of $V$, conditional on the same support, so 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 (maybe weakly) the bounds $\max\{s_i-1,\underline{V}\}$ and $\min\{s_i,\overline{V}\}$, which ensures that the Ineq. $\mathbb{E}[V|V+N=s_1] \geqslant \mathbb{E}[V|V+N=s_2]$ holds. ■

## A.6 Proof of Lemma 2

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

We begin by showing that $\mathbb{E}[V|A_1\cap A_2^{\mathrm{c}}] \geqslant \mathbb{E}[V|A_1^{\mathrm{c}}\cap A_2]$. 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

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

whereas

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

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

$$\begin{aligned}
\mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_2}(V+N_2)=1] &= \mathbb{E}[V|A_2]\\
&= \frac{p'}{p}\mathbb{E}\left[V|A_1\cap A_2\right]+\frac{p-p'}{p}\mathbb{E}\left[V|A_1^{\mathrm{c}}\cap A_2\right]\\
&\leqslant \frac{p'}{p}\mathbb{E}\left[V|A_1\cap A_2\right]+\frac{p-p'}{p}\mathbb{E}\left[V|A_1\cap A_2^{\mathrm{c}}\right]\\
&= \mathbb{E}[V|A_1]=\mathbb{E}[V|\hat{\sigma}_{\mathrm{SP}_1}(V+N_1)=1].
\end{aligned}$$

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

## A.7 Proof of Lemma 3

**Proof.** Assume, by contradiction, that there exists a random variable $N$, independent of $N_1$, such that $N_1+N\sim N_2\sim U[0,3/2]$. Evidently, $\mathrm{Supp}(N)\subseteq[0,1/2]$, otherwise $\mathrm{Supp}(N_1+N)\neq[0,3/2]$, 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

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

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