J. Econ. Theory 225 (2025) 105993

Contents lists available at ScienceDirect

Journal of Economic Theory

journal homepage: www.elsevier.com/locate/jet

Information orders in screening problems ☆

Andreas Asseyer

Freie Universität Berlin, Berlin School of Economics, Germany

ARTICLE INFO

JEL classification:
D82
D83

Keywords:
Screening problem
Information order
Mechanism design
Welfare analysis
Price discrimination

ABSTRACT

This paper studies information orders in screening models. I amend a general screening problem with a signal about the agent's type. The principal prefers one signal to another for any preferences of principal and agent if and only if the signals are ranked by Blackwell's order. Under a standard regularity condition, a novel information order – the hazard rate spread (HRS) order – characterizes a robust ranking of signals by the principal. I relate the HRS order to well-known information orders and provide sufficient conditions for other welfare measures than the principal's payoff to increase or decrease in the HRS order.

1. Introduction

In many markets, sellers conduct market research to gather information on the preferences of potential buyers. Market research reduces the informational advantage of buyers and enables the sellers to tailor their offers to the buyers' characteristics. But how should market research be conducted? Which information should a seller aim to gather? And how does market research affect the buyers and market efficiency?

I study these questions in a general model of monopolistic screening which I amend with a signal about the buyer's preferences. The signal realizes before the seller offers a mechanism to the buyer. Thus, the signal captures the information that the seller gathers through market research. By changing the signal, one can vary the information of the seller about the buyer's preferences and thereby analyze the effects of gradual changes in the information asymmetry between seller and buyer. Such an analysis complements the well-studied comparison of the extreme cases of the uninformed and the perfectly informed seller.

In this paper, I provide a robust comparison of signals about the preferences of a buyer with respect to the seller's profit and other welfare measures. As the first main result, I characterize the condition under which one signal is more profitable for a seller than another signal across a large class of preferences for the seller and the buyer. The second main result gives conditions under which a change to a seller-preferable signal increases or decreases welfare measures such as the buyer's surplus and efficiency. I use this result to generate new insights into the welfare effects of price discrimination with nonlinear tariffs in intermediate and final good markets. In particular, I show that total surplus is non-monotonic in the degree of information asymmetry between the seller and the buyer.

In the first part of the paper, I compare different signals from the seller's perspective. To this purpose, I analyze information orders for screening problems in the spirit of seminal work by Blackwell (1951, 1953). Blackwell studies when a decision maker prefers

☆ I thank Mikhail Drugov, Daniel Krähmer, Matthias Lang, Thomas Mariotti, Laurent Linnemer, Anne-Kathrin Roesler, Heiner Schumacher, Roland Strausz, three anonymous referees as well as audiences at Berlin, Bielefeld, Bonn, EARIE 2021, EEA-ESEM 2021, the French-German E-commerce Workshop, the meeting of the *Verein für Socialpolitik* 2021, and the meeting of the standing committee for industrial organization of the *Verein für Socialpolitik* 2024 for helpful comments and suggestions.

E-mail address: andreas.asseyer@fu-berlin.de.

https://doi.org/10.1016/j.jet.2025.105993
Received 17 April 2023; Received in revised form 4 March 2025; Accepted 5 March 2025

Available online 7 March 2025

one signal about an unobservable state over another for all possible preferences. He shows that such a robust ranking of signals is equivalent to one signal being a sufficient statistic for the joint observation of both signals.

As a benchmark result, I show that the characterization of Blackwell (1951, 1953) extends to screening problems if the set of admissible preferences is sufficiently rich. Theorem 1 states that the seller prefers one signal to another for any combination of quasilinear preferences for seller and buyer if and only if the signals are ranked in the Blackwell order. The implied equivalence between Blackwell's decision maker and the screening model hinges on preferences for which the seller cannot profitably screen the buyer so that all buyer types obtain the same allocation. The limited relevance of such preferences in applications motivates my subsequent focus on a smaller set of preferences.

The first main result of this paper identifies a novel information order that characterizes a robust ranking of signals by the seller for *regular* preferences of the seller and buyer.[1] The notion of *regularity* is a standard condition in the literature and ensures that screening is profitable.[2] Theorem 2 states that the seller prefers some signal over another for all *regular* preferences if and only if the signals are ranked by a novel informativeness criterion, the *hazard rate spread (HRS)*.

The HRS criterion ranks signals based on the distributions over inverse hazard rates which they induce. For a given type, different signal realizations induce different inverse hazard rates defined as the inverse probability ratios of the type to all higher types. Some signal is ranked ahead of another signal according to the HRS criterion if the distribution over inverse hazard rates induced by the first signal is a mean-preserving spread of the distribution induced by the other signal for any type.

Under regularity, the seller prefers signals that induce a more spread out distribution over inverse hazard rates. The seller's expected profit equals the expected virtual surplus, i.e., the difference between the total surplus and the buyer's expected information rent. Signals influence the virtual surplus only through the distribution over inverse hazard rates. With regularity, the optimal profit results from pointwise maximization of the virtual surplus which is linear in the inverse hazard rate. The resulting optimal virtual surplus is therefore convex in the inverse hazard rate – implying that the seller benefits from a more spread out distribution over inverse hazard rates.

I relate the HRS order to other information orders in the literature. The HRS order is weaker than Blackwell's order as the latter implies the first for any prior distribution over types while the reverse is not true (Proposition 2). For signals with the monotone likelihood ratio property, the HRS order is equivalent to the information orders of Lehmann (1988), Kim (1995), and Kim (2023) in the sense that two such signals are ranked by the HRS order for any prior if and only if the two signals are ranked by the equivalent orders of Lehmann (1988), Kim (1995), and Kim (2023) (Proposition 3). Finally, I show that the HRS order neither implies nor is implied by the information orders of Athey and Levin (2018) and Ganuza and Penalva (2010).

The second part of the paper demonstrates that the HRS order is a useful tool to compare signals with respect to other welfare measures than the seller's profit such as total or buyer's surplus. This part of the paper speaks to settings in which a regulator considers whether to interfere with the seller's efforts to gather information, e.g., through privacy regulation or restrictions on the use of information for price discrimination.

The main result of the second part provides simple sufficient conditions under which different welfare measures increase, respectively decrease, along the HRS order. For welfare measures that are linear in the payoffs of seller, buyer, and potentially affected third parties, the signal affects expected welfare only through the inverse hazard rates. As formally stated in Theorem 3, checking whether expected welfare is increasing or decreasing in the HRS order amounts to analyze for each type whether a univariate function derived from the payoff functions is concave or convex.

I apply the theorem to study the welfare effects of discriminatory nonlinear pricing on final and intermediate good markets. While the welfare comparison between a perfectly informed and an uninformed seller is clear-cut, I aim to shed light on the effects of a more gradual improvement of the seller's information. First, I provide conditions under which a better informed seller decreases total surplus in the classic model of Mussa and Rosen (1978). These conditions suggest that total surplus is non-monotone in the degree of information asymmetry between seller and buyer. Second, I uncover a tension between consumer surplus and total surplus in a model of nonlinear wholesale price discrimination along the lines of Herweg and Müller (2014). In particular, I consider a platform that intermediates the trade between the seller and the buyer. If the platform becomes better informed about the production costs of the sellers, total surplus may decrease but consumers still benefit. This result shows that the objective of competition authorities may critically affect their regulatory policies on intermediated markets.

In the remainder of this section, I discuss the related literature. Section 2 introduces the model. Section 3 characterizes robust comparisons of signals from the seller's perspective. Section 4 studies the welfare implications of changes to seller-preferable signals. Section 5 concludes. Omitted proofs are collected in Appendix A.

1.1. Relation to the literature

This paper contributes to the literature on price discrimination (Pigou, 1920; Robinson, 1933; Schmalensee, 1981). Aguirre et al. (2010) consider a market consisting of two market segments. They provide conditions such that price discrimination between the two market segments increases or decreases welfare in comparison to the case where the monopolist has to set the same price for the whole market. Their approach can therefore be reframed as the comparison of an uninformative signal with a binary signal which

[1] In analogy to the literature which analyzes information orders for restricted subsets of decision problems (Lehmann, 1988; Athey and Levin, 2018; Kim, 2023), I focus attention on screening problems in which the principal and the agent exhibit regular preferences.

[2] As further defined below, regularity requires that the seller's problem can be solved using the Myersonian – or first-order – approach.

splits the market into the two segments. Bergemann et al. (2015) take a different approach. They characterize the set of payoffs for seller and buyer that can arise for some arbitrary market segmentation, i.e., an arbitrary signal on the buyer's preferences. As the vast majority of the literature on price discrimination, these papers restrict attention to linear prices. Two notable exceptions are Herweg and Müller (2014), who study nonlinear price discrimination on intermediate good markets, and Haghpanah and Siegel (2023) who analyze nonlinear price discrimination on final good markets. Herweg and Müller (2014) follow again the approach of Aguirre et al. (2010) and compare total surplus under a ban on price discrimination – equivalent to an uninformative signal – with the total surplus obtained under a binary market segmentation. Haghpanah and Siegel (2023) follow the approach of Bergemann et al. (2015) and show that any generic market, or market segment, can be further segmented to obtain a Pareto improvement for the seller and the buyer for all types. Thus, the seller's information can in their framework always be refined such that total surplus increases. The current paper aims to combine these approaches. Like Aguirre et al. (2010), I compare two signals, which speaks to the fact that in practice only a restricted subset of theoretically plausible signals may be available. Like Bergemann et al. (2015), I allow for a large class of possible signals. In Section 4.2, I show how my approach generates novel results on the effects of nonlinear price discrimination in intermediate and final good markets.

This paper is furthermore related to the literature on informativeness criteria in principal-agent models. Most of this literature has hitherto focused on the case of moral hazard. Holmström (1979) shows that the principal benefits from conditioning the agent's incentive contract on an additional signal if and only if the signal is informative about the agent's action. Comparing two signals about the agent's action, Gjesdal (1982) and Grossman and Hart (1983) show that a Blackwell more informative signal is weakly preferred by the principal independently of the specific preferences of principal and agent. For agency models in which the first-order approach is valid, Kim (1995) provides a weaker sufficient condition for the principal to prefer one signal over another. This condition requires that the distribution over the likelihood ratios induced by the preferred signal is a mean-preserving spread of the distribution induced by the other signal. Jewitt (2007) proves that the order of Kim (1995) is also necessary for a robust signal ranking under validity of the first-order approach. This paper's HRS order can be seen as the analogue of Kim (1995)'s condition for the context of screening problems.[3]

Further related is the literature on information orders with a single decision maker. While Blackwell (1951, 1953) provides an information order for the full domain of decision problems, Lehmann (1988), Quah and Strulovici (2009), Di Tillio et al. (2021), and Kim (2023) characterize information orders for different domains of monotone decision problems.[4] Lehmann (1988), Quah and Strulovici (2009), and Di Tillio et al. (2021) consider settings in which the decision maker chooses an action from a one-dimensional set. This is unlike a principal in a monopolistic screening model who chooses a mechanism which is a higher-dimensional object. In a recent closely related paper, Kim (2023) provides a novel information order for a setting with a single decision maker who faces a multidimensional decision problem. Kim uses this information order to derive sufficient conditions for a principal to prefer one signal over another in a screening problem. The current paper is complementary to the contribution of Kim (2023) in three ways. First, the HRS order allows the ranking of additional signals that do not fall into the domain considered by Kim (2023). Second, my results can be used to show that the information order of Kim (2023) is also a necessary condition for a robust comparison of signals from the principal's perspective. Third, the HRS order is a helpful tool to extend the comparison of signals to other welfare measures than the principal's profit. In Section 3.3.3, I discuss the relation between the two papers in more detail.

This paper furthermore contributes to the literature on robustness in screening problems. An important strand of this literature studies the set of outcomes that can be predicted knowing the players' payoff functions but not the information structure (Bergemann and Morris, 2013, 2016; Bergemann et al., 2015, 2017). Another strand assumes knowledge of the information structure but not of payoff functions. In several of these contributions, it is the principal who lacks knowledge on payoff functions and evaluates mechanisms by their worst case outcome (Chassang, 2013; Garrett, 2014; Carroll, 2015). By contrast, Garrett (2021) considers an observer outside of the principal-agent relationship. For a given information structure, he provides a robust bound on the ratio of the principal's to the agent's payoff. The current paper contributes to this literature by providing a robust ranking of signals with respect to the principal's payoff and other welfare measures.

The following papers also amend a screening model with an ex-ante signal.[5] In closely related work, Ottaviani and Prat (2001) consider a monopolistic screening model in which the agent has imperfect information about his preferences and the principal can generate an informative public signal. They provide conditions on the agent's preferences and the information structure under which the principal benefits from fully releasing the public signal.[6] The current paper assumes the agent to have perfect information about preferences, which is a special case of the framework in Ottaviani and Prat (2001). At the same time, the current paper compares a larger class of signals in that it does not require the signals to be Blackwell ordered. Using the model of Laffont and Tirole (1986), Boyer and Laffont (2003) analyze whether the additional signal strengthens incentives. Drugov (2010) studies a model where the principal may delay contracting to wait for an additional signal and shows that social surplus may decrease in the quality of the signal.[7]

[3] Dewatripont et al. (1999) study the effects of the information structure on incentives in the career concerns model.

[4] The approach of this literature to restrict the domain of decision problems to rank additional signals is also mirrored in Theorems 1 and 2 of this paper. The restriction to regular screening problems can therefore be seen in analogy to the restriction to monotone decision problems with a single decision maker.

[5] An ex-ante signal realizes before the principal offers a mechanism. Another strand of literature studies the effects of a signal that arrives after contracting (Riordan and Sappington, 1988; Demougin and Garvie, 1991; Gary-Bobo and Spiegel, 2006; Bose and Zhao, 2007).

[6] Saak (2007) further weakens the conditions of Ottaviani and Prat (2001). See Cella (2008) for a setting in which the principal may benefit from keeping a signal private.

[7] Unpublished work by Drugov (2011) focuses on the second aspect in a screening model with two possible types.

Finally, I would like to point out that the results of this paper may be insightful for other applications of the monopolistic screening model than the nonlinear pricing of products (Mussa and Rosen, 1978; Maskin and Riley, 1984), such as the regulation of monopolies (Baron and Myerson, 1982), the procurement of intermediate goods (Laffont and Tirole, 1986), the allocation of scarce resources within firms (Harris et al., 1982), or the taxation of labor income (Diamond, 1998).

2. Model

I amend a screening model under quasilinear preferences with a signal that reflects how well the seller (henceforth, the principal) is informed about the preferences of the buyer (the agent).

2.1. Screening model

There is a principal and an agent. The agent privately observes the type θ from the finite set $\Theta \equiv \{\theta_1, \dots, \theta_m\}$. The type is drawn from a commonly known prior distribution with the strictly positive probability mass function $f(\theta_i)$. I denote the cumulative distribution function by $F(\theta_i) \equiv \sum_{k=1}^{i} f(\theta_k)$. The preferences of principal and agent over the decision $x \in X \subseteq \mathbb{R}^k$, with X being compact, and the transfer $t \in \mathbb{R}$ depend on the agent's type, and are quasilinear in the transfer. For a type $\theta \in \Theta$ and a decision $x \in X$, the gross payoff of the principal is $u^P(\theta, x)$ and the gross payoff of the agent is $u^A(\theta, x)$. For each $\theta \in \Theta$, the functions $u^A(\theta, \cdot)$ and $u^P(\theta, \cdot)$ are bounded and continuous in X. The (net) payoffs of principal and agent for the decision $x \in X$ and the transfer $t \in \mathbb{R}$ are therefore given by

$$u^P(\theta, x) + t \quad \text{and} \quad u^A(\theta, x) - t.$$

A stochastic decision $q \in Q \equiv \Delta X$ is a probability distribution over the set X. The (expected) gross payoffs from a stochastic decision $q \in Q$ are

$$v^P(\theta, q) \equiv \int_X u^P(\theta, x) dq(x) \quad \text{and} \quad v^A(\theta, q) \equiv \int_X u^A(\theta, x) dq(x).$$

Due to the quasilinearity of preferences, it is without loss of generality to neglect stochastic transfers between the two parties. The agent has an outside option $O \in X$. The agent's value for the outside option is normalized to zero, i.e., $u^A(\theta, O) = 0$ for all $\theta \in \Theta$.[8] I denote the joint surplus of principal and agent by

$$\bar{u}(\theta, x) \equiv u^P(\theta, x) + u^A(\theta, x) \quad \text{and} \quad \bar{v}(\theta, q) \equiv v^P(\theta, q) + v^A(\theta, q).$$

The set of all triples (X, u^P, u^A) that satisfy the conditions above is denoted by $\mathcal{U}$.

2.2. Signals

A signal determines the information asymmetry between principal and agent without changing the underlying prior distribution of the agent's type. Following Blackwell (1951), a signal $\sigma = (S, \{g(\cdot|\theta)\}_\Theta)$ consists of a finite set $S = \{s_1, \dots, s_n\}$ of signal realizations and a conditional probability mass function $g(s|\theta)$ over S for each type $\theta \in \Theta$. Without loss of generality, let S be the union of the supports of the probability mass functions in $\{g(\cdot|\theta)\}_\Theta$. The signal realization is observed by the principal and the agent. For the prior f and the signal σ, the joint distribution of signal and type, the marginal distribution over signals, and the conditional distribution over types given a signal realization are

$$g_\sigma^f(\theta, s) \equiv g(s|\theta) f(\theta), \quad g_\sigma^f(s) \equiv \sum_{i=1}^{m} g_\sigma^f(\theta_i, s), \quad \text{and} \quad g_\sigma^f(\theta|s) \equiv \frac{g_\sigma^f(\theta, s)}{g_\sigma^f(s)}.$$

The associated cumulative distribution functions are denoted by capital letters. Let $\Theta_\sigma(s) \equiv \{\theta \in \Theta : g(s|\theta) > 0\}$ be the conditional support of the type for the signal realization $s \in S$. Last, I define the inverse hazard rate of the posterior distribution upon receiving the signal realization s by $h_\sigma^f(\cdot|s) : \Theta_\sigma(s) \to \mathbb{R}_+$ with

$$h_\sigma^f(\theta|s) \equiv \frac{1 - G_\sigma^f(\theta|s)}{g_\sigma^f(\theta|s)}.$$

2.3. Mechanisms

The principal offers the agent a mechanism after having observed a realization $s \in S$ of the signal σ. Due to the revelation principle, the principal cannot do better than to offer a direct mechanism which incentivizes the agent to participate and to report the type truthfully. Formally, a direct mechanism $(q(\cdot|s), t(\cdot|s))$ consists of a decision rule $q(\cdot|s) : \Theta_\sigma(s) \to Q$ and a transfer rule $t(\cdot|s)$:

[8] Such a normalization is equivalent to a linear transformation $u^A(\theta, \cdot) \equiv \tilde{u}^A(\theta, \cdot) - \tilde{u}^A(\theta, O)$ of an initial payoff function $\tilde{u}^A$.

$\Theta_\sigma(s) \to \mathbb{R}$, which determine an allocation (q,t) as functions of the self-reported type $\theta' \in \Theta_\sigma(s)$. A direct mechanism incentivizes the agent to report truthfully if it satisfies the incentive-compatibility constraint

$$v^A(\theta, q(\theta|s)) - t(\theta|s) \geq v^A(\theta, q(\theta'|s)) - t(\theta'|s), \quad \forall \theta, \theta' \in \Theta_\sigma(s). \tag{$IC_{\sigma:s}(X,u^A)$}$$

The agent is willing to accept such a mechanism if it satisfies the participation constraint[9]

$$v^A(\theta, q(\theta|s)) - t(\theta|s) \geq 0, \quad \forall \theta \in \Theta_\sigma(s). \tag{$PC_{\sigma:s}(X,u^A)$}$$

A direct mechanism $(q(\cdot|s), t(\cdot|s))$ is feasible for the realization $s \in S$ of the signal σ if it satisfies ($IC_{\sigma:s}(X,u^A)$) and ($PC_{\sigma:s}(X,u^A)$). A collection of mechanisms $\{(q(\cdot|s), t(\cdot|s))\}_{s\in S}$ is feasible for the signal σ if $(q(\cdot|s), t(\cdot|s))$ is feasible for any signal realization $s \in S$. I denote the set of feasible collections of mechanisms for the signal σ by $\mathcal{M}_\sigma(X,u^A)$.[10] The principal's optimal choice of a feasible collection of mechanisms for the signal σ, the prior f and the preferences (X, u^P, u^A) solves the optimization problem

$$\max_{\mathcal{M}_\sigma(X,u^A)} \sum_{i=1}^{m} \sum_{j=1}^{n} \left(v^P(\theta_i, q(\theta_i|s_j)) + t(\theta_i|s_j)\right) g(s_j|\theta_i) f(\theta_i). \tag{$\mathbf{P}^f_\sigma(X,u^P,u^A)$}$$

The principal's optimal payoff in the optimization problem ($\mathbf{P}^f_\sigma(X,u^P,u^A)$) is denoted by $\Pi^f_\sigma(X,u^P,u^A)$.

In Proposition 3, I consider the situation where the principal is restricted to offer deterministic direct mechanisms. Formally, the principal then faces the problem ($\mathbf{P}^f_\sigma(X,u^P,u^A)$) with the additional constraint that for all θ and s, $q(\theta|s) = \delta_{x(\theta|s)}$ with $x(\theta|s) \in X$, where δ_x is the Dirac measure on $x \in X$. I denote the principal's optimal payoff under this optimization problem by $\underline{\Pi}^f_\sigma(X,u^P,u^A)$.

2.4. Timing

The interaction between principal and agent unfolds as follows. First, nature draws the agent's type θ and the signal realization s. Second, the agent observes θ and s whereas the principal observes only s. Third, the principal offers the agent a direct mechanism. Finally, the agent decides whether to accept or reject the mechanism. If the agent rejects, the outside option O is implemented. If the agent accepts, the mechanism is played.

2.5. Example

I use the following running example in the spirit of Mussa and Rosen (1978) to illustrate my results. The principal is a seller and the agent is a buyer. The principal can produce $x \in [0,1]$ units of a good at cost $0.5x^2$. The agent values x units by θx where θ is uniformly drawn from $\{\theta_1, \theta_2\} \subset \mathbb{R}$ with $0 < \theta_1 < \theta_2 \leq 1$ and $\Delta\theta \equiv \theta_2 - \theta_1$. A binary signal σ generates signal realizations s_1 and s_2 with $\Pr(s_i|\theta_i) = \rho \in [0.5, 1)$ for $i = 1,2$. Hence, the signal realization s_i induces the posterior belief $\Pr(\theta_i|s_i) = \rho$ for $i = 1,2$. The principal offers the agent a mechanism in the form of an incentive-compatible menu of price-quantity pairs $\{(t_1, x_1), (t_2, x_2)\}$.[11] Thus, the type θ_i prefers the contract (t_i, x_i) over the other contract and the outside option of not buying, i.e., $\theta_i x_i - t_i \geq \theta_i x_j - t_j$ and $\theta_i x_i - t_i \geq 0$ for $i,j = 1,2$.

3. Robust signal comparisons

Throughout this section, I compare two signals $\check{\sigma} = (\check{S}, \{\check{g}(\cdot|\theta)\}_{\theta\in\Theta})$ with $\check{S} = \{\check{s}_1, \dots, \check{s}_{\check{n}}\}$ and $\hat{\sigma} = (\hat{S}, \{\hat{g}(\cdot|\theta)\}_{\theta\in\Theta})$ with $\hat{S} = \{\hat{s}_1, \dots, \hat{s}_{\hat{n}}\}$ on different domains of preferences.

3.1. Benchmark: full preference domain

In this section, I provide as a benchmark a robust comparison of signals from the principal's perspective for the full domain of preferences. In particular, I seek to characterize the relationship between two signals $\check{\sigma}$ and $\hat{\sigma}$ for which the expected payoff of the principal is higher under $\check{\sigma}$ than under $\hat{\sigma}$ for any combination of preferences $(X, u^P, u^A) \in \mathcal{U}$.

First, I recall the information order due to Blackwell (1951, 1953).

Definition 1. The signal $\check{\sigma}$ is ranked ahead of the signal $\hat{\sigma}$ in the Blackwell order, denoted by $\check{\sigma} \succeq_B \hat{\sigma}$, if there exists a matrix $B = [b_{jk}] \in \mathbb{R}^{\check{n}\times\hat{n}}$ such that

1. $b_{jk} \geq 0, \quad \forall j \in \{1, \dots \check{n}\}, k \in \{1, \dots, \hat{n}\}$,
2. $\sum_{k=1}^{\hat{n}} b_{jk} = 1, \quad \forall j \in \{1, \dots, \check{n}\}$,

[9] Note that there is no loss in inducing the agent to always participate in the mechanism as $O \in X$.

[10] Note that the set $\mathcal{M}_\sigma(X,u^A)$ depends on the signal through the support of the posterior beliefs induced by the signal realizations. $\mathcal{M}_\sigma(X,u^A)$ does not depend on the prior.

[11] This is without loss of optimality.

3. $\hat{g}(\hat{s}_k|\theta_i) = \sum_{j=1}^{\check{n}} b_{jk}\check{g}(s_j|\theta_i), \quad \forall i = \{1,\ldots,m\}, k \in \{1,\ldots,\hat{n}\}$.

A signal $\check{\sigma}$ is ranked ahead of signal $\hat{\sigma}$ in the Blackwell order if observing $\hat{\sigma}$ is informationally equivalent to observing $\check{\sigma}$ with some noise, i.e., the distribution over posterior beliefs are identical in the two instances. Blackwell (1951, 1953) shows that a decision maker who has to make a choice $x \in X$ under uncertainty about a state $\theta \in \Theta$ prefers some signal $\check{\sigma}$ over another signal $\hat{\sigma}$ for any payoff function $u^D : \Theta \times X \to \mathbb{R}$ if and only if $\check{\sigma}$ is ranked ahead of $\hat{\sigma}$ in Blackwell's order. I obtain a closely related result in the screening model for the full domain of preferences.

Theorem 1. *Given a prior f and two signals $\check{\sigma}$ and $\hat{\sigma}$,*

$$\Pi^f_{\check{\sigma}}(X,u^P,u^A) \geq \Pi^f_{\hat{\sigma}}(X,u^P,u^A) \ \forall (X,u^P,u^A) \in \mathcal{U} \iff \check{\sigma} \succeq_B \hat{\sigma}.$$

The principal prefers one signal to another for all possible preference profiles in the set $\mathcal{U}$ if and only if the first signal Blackwell dominates the second. Thus, the theorem also proves the equivalence of a robust comparison of a principal who can consult an agent who knows the state θ with that of a decision maker who has to make a choice without further information than that provided by the signal.

The sufficiency of Blackwell's ordering is intuitive. Like Blackwell's decision maker, the principal cannot be made worse off by a less noisy signal as it is always possible to ignore the additional information. In particular, for any given signal $\check{\sigma}$, the principal can replicate the outcome obtained under a signal $\hat{\sigma} \preceq_B \check{\sigma}$ by mixing over the mechanisms used under $\hat{\sigma}$ with probabilities derived from the matrix B of Definition 1.

I prove the necessity of Blackwell's ordering by using preferences for principal and agent that are fully misaligned. Under these *adversarial* preferences, the principal does not benefit from screening the agent and – like Blackwell's decision maker – has to make a decision just based on the signal realization. Take some payoff function $u^D(\theta,x)$ for Blackwell's decision maker and assign the principal and the agent the payoff functions

$$u^P(\theta,x) = (1+\alpha)u^D(\theta,x) \quad \text{and} \quad u^A(\theta,x) = -\alpha u^D(\theta,x)$$

for some $\alpha > 0$. Under these preferences, the principal does not benefit from screening the agent. Moreover, as α converges to zero, the principal's optimal expected payoff converges to the optimal payoff of the decision maker. Thus, the necessity of Blackwell's order follows from Blackwell's theorem.

Theorem 1 extends to richer domains of preferences that include but are not restricted to quasilinear preferences. This follows from the observation that the linearity of payoffs in the transfer is not used in the proof of Lemma 3 in Appendix A. The restriction to quasilinear preferences alone is not enough to obtain a robust signal ranking that is weaker than the demanding criterion of Blackwell.[12]

Adversarial preferences are very special in that they strongly restrict the set of decision rules which the principal can implement. In particular, the first-best decision rule is typically not even implementable under these preferences. This raises the question whether a weaker information order could be obtained if one were to focus on a smaller set of admissible preferences which excludes adversarial preferences.

3.2. Main result: regular preference domain

This section presents the main result of the paper. I provide a robust comparison of two signals on the domain of preferences that satisfy a regularity condition. As the previous section suggests, the equivalence between the robust signal comparison on the full preference domain and Blackwell's order relies on preferences under which the principal cannot profitably screen the agent. Such preferences are certainly special and play a limited role in applications of the screening model. Thus, one might wonder how a robust ranking of signals could be characterized if screening is profitable. In this section, I use a regularity condition that is widely used in the literature and which ensures that screening is profitable.

The regularity condition requires that the principal's problem ($\mathbf{P}^f_\sigma(X,u^P,u^A)$) can be solved using the *Myersonian approach* (Myerson, 1981; Baron and Myerson, 1982).[13] It is the standard solution approach to the principal's problem and widely used in applications. The approach consists in solving a relaxed version of problem ($\mathbf{P}^f_\sigma(X,u^P,u^A)$) and imposing conditions under which the solution to the relaxed problem also solves the original problem. These conditions are commonly referred to as regularity conditions. In particular, one relaxes problem ($\mathbf{P}^f_\sigma(X,u^P,u^A)$) by neglecting for any $s \in S$ all constraints but the participation constraint of the lowest type and the local downward incentive-compatibility constraints of all higher types.

To keep the identity of the lowest type as well as the relationship of adjacency among types stable, I focus on signals that satisfy a full support condition. A signal has full support if the principal can never rule out any type of the agent after observing the signal realization. Formally, this is defined as follows.

Definition 2. A signal σ has full support if $\Theta_\sigma(s) = \Theta$ for all $s \in S$.

[12] By contrast, if u^P (u^A) was fixed, a weaker order could characterize the principal's preference for one signal for all u^A (u^P).

[13] The literature alternatively refers to this as first-order or local approach.

For a given signal with full support, one can then formulate the relaxed problem of the principal as follows. For each signal realization $s \in S$, the participation constraint of the lowest type is

$$v^A(\theta_1, q(\theta_1|s)) - t(\theta_1|s) \geq 0. \tag{$RPC_{\sigma:s}(X,u^A)$}$$

The local downward incentive-compatibility constraints for $s \in S$ are

$$v^A(\theta_i, q(\theta_i|s)) - t(\theta_i|s) \geq v^A(\theta_i, q(\theta_{i-1}|s)) - t(\theta_{i-1}|s) \;\; \forall i = 2, \dots, m. \tag{$RIC_{\sigma:s}(X,u^A)$}$$

Let the set $\mathcal{R}_\sigma(X, u^A)$ consist of all collections of mechanisms that satisfy ($RPC_{\sigma:s}(X,u^A)$) and ($RIC_{\sigma:s}(X,u^A)$) for the signal σ. The relaxed problem is then given by

$$\max_{\mathcal{R}_\sigma(X,u^A)} \sum_{i=1}^{m} \sum_{j=1}^{n} \left(v^P(\theta_i, q(\theta_i|s_j)) + t(\theta_i|s_j) \right) g(s_j|\theta_i) f(\theta_i). \tag{$\mathbf{R}^f_\sigma(X,u^P,u^A)$}$$

The relaxed problem ($\mathbf{R}^f_\sigma(X,u^P,u^A)$) is simpler to solve than the original problem ($\mathbf{P}^f_\sigma(X,u^P,u^A)$). It is straightforward to show that the constraints ($RPC_{\sigma:s}(X,u^A)$) and ($RIC_{\sigma:s}(X,u^A)$) are always satisfied with equality in the optimum. The binding constraints pin down the transfer rule as a function of the decision rule. Substituting into the objective function, basic manipulations yield the unconstrained objective

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \left(\bar{v}(\theta_i, q(\theta_i|s_j)) - h^f_\sigma(\theta_i|s_j) \Delta^A_v(\theta_i, q(\theta_i|s_j)) \right) g(s_j|\theta_i) f(\theta_i)$$

with $\Delta^A_v(\theta_i, q) \equiv v^A(\theta_{i+1}, q) - v^A(\theta_i, q)$ for $i = 1, \dots, m-1$ and $\Delta^A_v(\theta_m, \cdot) = 0$. The function $\Delta^A_v(\theta_i, q)$ measures the incremental rent that has to be given to types $\theta > \theta_i$ to disincentivize the type θ_{i+1} to falsely report θ_i when this report induces the decision q. The unconstrained objective is maximal for the decision that maximizes for each type θ_i the *virtual joint surplus*

$$\bar{v}(\theta_i, q) - h^f_\sigma(\theta_i|s_j) \Delta^A_v(\theta_i, q),$$

where the inverse hazard rate $h^f_\sigma(\theta_i|s_j)$ reflects the relative frequency of the type θ_i – for whom the joint surplus $\bar{v}(\theta_i, q)$ accrues – to all higher types – whose information rent contains $\Delta^A_v(\theta_i, q)$.

Under regularity, the solution to the relaxed problem is also a solution to the original problem. One may impose assumptions on the payoff functions and the type distributions such that the solution to the relaxed problem satisfies the neglected constraints. Here, I do not invoke such assumptions but instead work directly with the following condition.[14]

Definition 3. The triple $(X, u^P, u^A) \in \mathcal{U}$ is regular for the prior f and the signal σ if there exists a solution to ($\mathbf{R}^f_\sigma(X,u^P,u^A)$) that solves ($\mathbf{P}^f_\sigma(X,u^P,u^A)$). The set of all such triples is denoted by $\mathcal{U}^{R,f}_\sigma$.

As an important step toward the main result of this section, I show that – given full support and regularity – the principal's payoff depends on the signal and its realization only through the impact on the inverse hazard rates. With full support, the higher adjacent type of some type does not change with the signal or the signal realization. It follows that the maximal virtual joint surplus can be expressed as a function of the type and the value of the inverse hazard rate only. For the type $\theta_i \in \Theta$ and the value of the inverse hazard rate z, the maximum of the virtual joint surplus is

$$\max_{Q} \bar{v}(\theta_i, q) - z \Delta^A_v(\theta_i, q).$$

Under regularity, the principal's optimal ex-ante payoff is the expected maximal virtual joint surplus. As the signal affects the maximal virtual surplus only through the value of the inverse hazard rate under full support, the ex-ante payoff is affected by the signal only through the distribution over the values of the inverse hazard rates. For a given prior f and type $\theta \in \Theta$, the signal σ induces a random variable that takes values in the set $\{h^f_\sigma(\theta|s)\}_{s \in S}$ according to the cumulative distribution function[15]

$$M^f_\sigma(z|\theta) \equiv \sum_{j=1}^{n} g(s_j|\theta) \mathbf{1}(h^f_\sigma(\theta|s_j) \leq z).$$

Under regularity and full support, the signal affects the principal's ex-ante payoff only through the distribution $M^f_\sigma(z|\theta)$. Thus, the principal's optimal ex-ante payoff is the expectation of the maximal virtual surplus over the type and the inverse hazard rate.

[14] This approach is also used by Dequiedt and Martimort (2015). I provide sufficient conditions for regularity in Proposition 1.

[15] The indicator function $\mathbf{1}(A)$ takes the value 1 if A holds and zero otherwise.

Lemma 1. *Suppose the signal σ has full support. Then, $(X, u^P, u^A) \in \mathcal{U}_\sigma^{R,f}$ implies*

$$\Pi_\sigma^f(X, u^P, u^A) = \sum_{i=1}^{m} \int_0^\infty \left\{ \max_Q \bar{v}(\theta_i, q) - z\Delta_v^A(\theta_i, q) \right\} dM_\sigma^f(z|\theta_i) f(\theta_i).$$

The lemma suggests that signals may be ordered based on the distributions over inverse hazard rates they induce. I propose the following information order.

Definition 4. For a given prior f, the signal $\check{\sigma}$ is ranked ahead of the signal $\hat{\sigma}$ in the Hazard Rate Spread (HRS) order, denoted by $\check{\sigma} \succeq_{HRS}^f \hat{\sigma}$, if for all $\theta \in \Theta$ and $z \geq 0$

$$\int_0^z M_{\check{\sigma}}^f(y|\theta) dy \geq \int_0^z \hat{M}_{\hat{\sigma}}^f(y|\theta) dy.$$

A signal $\check{\sigma}$ dominates a signal $\hat{\sigma}$ in the HRS order if the distribution over inverse hazard rates induced by $\check{\sigma}$ is a mean-preserving spread of the distribution induced by $\hat{\sigma}$ for any type $\theta \in \Theta$.[16] The following result shows that the HRS order characterizes a robust ranking of signals for the principal in regular environments.

Theorem 2. *Given a prior f and two signals $\check{\sigma}$ and $\hat{\sigma}$ with full support,*

$$\Pi_{\check{\sigma}}^f(X, u^P, v^A) \geq \Pi_{\hat{\sigma}}^f(X, u^P, u^A) \;\; \forall (X, u^P, u^A) \in \mathcal{U}_{\check{\sigma}}^{R,f} \iff \check{\sigma} \succeq_{HRS}^f \hat{\sigma}.$$

Under regularity, the principal robustly ranks some signal $\check{\sigma}$ over another signal $\hat{\sigma}$ if and only if $\check{\sigma}$ is ranked ahead of $\hat{\sigma}$ in the HRS order.

First, I argue that the HRS order is a sufficient condition for the principal to prefer one signal over another for all regular preference combinations. Consider the formulation of the principal's expected payoff in Lemma 1. Note that the linearity of the virtual surplus in the inverse hazard rate z implies – by a standard revealed preference argument – that the maximal virtual surplus $\max_Q \bar{v}(\theta_i, q) - z\Delta_v^A(\theta_i, q)$ is convex in z. Due to Jensen's inequality, the principal prefers more spread out distributions over inverse hazard rates just as a risk-loving individual would prefer a riskier lottery over outcomes. Hence, the principal prefers the signal that is ranked higher according to the HRS order.

It remains to show that the hazard rate spread criterion is a necessary condition for a robust signal ranking of the principal. Take two signals $\check{\sigma} \not\succeq_{HRS}^f \hat{\sigma}$. Then there exist $\tilde{\theta} \in \Theta$ and $\tilde{z} > 0$ such that

$$\int_0^{\tilde{z}} M_{\check{\sigma}}^f(z|\tilde{\theta}) dz < \int_0^{\tilde{z}} M_{\hat{\sigma}}^f(z|\tilde{\theta}) dz.$$

I argue that it is possible to construct preferences under which the principal obtains a higher payoff under $\hat{\sigma}$. Consider a bilateral trade environment with private values given by $\tilde{X} = \{0, 1\}$, $\tilde{u}^P(\theta_i, x) = 0$, and $\tilde{u}^A(\theta_i, x) = v(\theta_i)x$. For all types $\theta_i > \tilde{\theta}$, set $v(\theta_i) = 1$. Assign all types $\theta_i < \tilde{\theta}$ a negative valuation. Finally, set $v(\tilde{\theta}) = \tilde{z}/(1 + \tilde{z})$. These preferences replicate a model with two types as the types below $\tilde{\theta}$ can be excluded at no cost. Thus, the preferences are regular for any f and σ. Using Lemma 1, the principal's expected payoff under a generic signal σ can be expressed as

$$1 - F(\tilde{\theta}) + f(\tilde{\theta}) \int_0^\infty \max_{q \in [0,1]} q \left(\frac{\tilde{z}}{1 + \tilde{z}} - z \left(1 - \frac{\tilde{z}}{1 + \tilde{z}} \right) \right) dM_\sigma^f(z|\tilde{\theta})$$

$$= 1 - F(\tilde{\theta}) + f(\tilde{\theta}) \frac{\int_0^{\tilde{z}} M_\sigma^f(z|\tilde{\theta}) dz}{1 + \tilde{z}}.$$

Thus, the expected payoff of the principal is higher under $\hat{\sigma}$ than under $\check{\sigma}$ for these preferences.

Theorem 2 relies neither on preferences that are nonlinear in the decision x nor on common values. Indeed, as the proof of necessity shows, the HRS order characterizes the robust signal ranking even in regular models of bilateral trade with private values.

[16] With full support, second-order stochastic dominance coincides with mean-preserving spread as $\int_0^\infty z dM_\sigma^f(z|\theta_i) = \sum_{j=1}^n h_\sigma^f(\theta_i|s_j) g(s_j|\theta_i) = \sum_{j=1}^n h_\sigma^f(\theta_i|s_j) \frac{g_\sigma^f(\theta_i|s_j) g_\sigma^f(s_j)}{f(\theta_i)} = \frac{\sum_{j=1}^n (1 - G_\sigma^f(\theta_i|s_j)) g_\sigma^f(s_j)}{f(\theta_i)} = \frac{1 - F(\theta_i)}{f(\theta_i)}$.

3.3. Discussion of the main result

In this section, I first illustrate the sufficiency of the HRS order in the context of the simple example presented in Section 2.5. I then provide sufficient conditions for a screening problem to be regular for a given signal. Finally, I relate the HRS order to other information orders.

3.3.1. Sufficiency of HRS order in the example

Recall the simple version of the model from Section 2.5. I show that an increase in the precision parameter ρ benefits the principal as it induces a more spread out distribution of inverse hazard rates. For each signal realization $s_i \in \{s_1, s_2\}$, the principal aims to maximize her expected profit. Under any optimal menu of price-quantity pairs, the participation constraint of type θ_1 and the incentive constraint of type θ_2 are binding. This yields the transfers $t_1 = \theta_1 x_1$ and $t_2 = \theta_2 x_2 - \Delta\theta x_1$. Substituting these into the principal's objective, we obtain the following expression for the expected profit:

$$\Pr(\theta_1|s_i)\left(\theta_1 x_1 - 0.5x_1^2 - \frac{1-\Pr(\theta_1|s_i)}{\Pr(\theta_1|s_i)}\Delta\theta x_1\right) + (1-\Pr(\theta_1|s_i))(\theta_2 x_2 - 0.5x_2^2).$$

The profit-maximizing quantity for the high type θ_2 is $x_2 = \theta_2$ and therefore independent of the signal realization. The profit-maximizing quantity for the low type depends on the signal realization through the inverse hazard rate $\frac{1-\Pr(\theta_1|s_i)}{\Pr(\theta_1|s_i)}$ and is given by $x_1 = [\theta_1 - \frac{1-\Pr(\theta_1|s_i)}{\Pr(\theta_1|s_i)}\Delta\theta]_+$.[17]

We can now express the ex-ante expected payoff as a function of the distribution over inverse hazard rates as in Lemma 1. The distribution $M_\sigma^f(z|\theta_1)$ over inverse hazard rates conditional on $\theta = \theta_1$ is binary. It takes the value $\frac{1-\rho}{\rho}$ with probability ρ and the value $\frac{\rho}{1-\rho}$ with probability $1-\rho$.[18] The ex-ante expected payoff is

$$0.25\left(\rho\left[\theta_1 - \frac{1-\rho}{\rho}\Delta\theta\right]_+^2 + (1-\rho)\left[\theta_1 - \frac{\rho}{1-\rho}\Delta\theta\right]_+^2\right) + 0.25\theta_2^2.$$

An increase in the precision ρ leads to a more spread out distribution over inverse hazard rates as $\frac{1-\rho}{\rho}$ decreases while $\frac{\rho}{1-\rho}$ increases. To see that this benefits the principal, note that the expression $\left[\theta_1 - z\Delta\theta\right]_+^2$ is convex in z. Thus, a more spread out distribution over inverse hazard rates increases the ex-ante expected payoff by Jensen's inequality. As an example, one can compare the ex-ante expected payoff of an informative signal, i.e., $\rho > 0.5$ and an uninformative signal with $\rho' = 0.5$. Assuming that production is strictly positive for the type θ_1 under all signal realizations, this yields a strictly positive payoff difference of

$$\begin{aligned}0.25&\left(\rho\left(\theta_1 - \frac{1-\rho}{\rho}\Delta\theta\right)^2 + (1-\rho)\left(\theta_1 - \frac{\rho}{1-\rho}\Delta\theta\right)^2\right) - 0.25\left(\theta_1 - \Delta\theta\right)^2 \\ &= 0.25\Delta\theta^2\left(\frac{(2\rho-1)^2}{(1-\rho)\rho}\right) > 0.\end{aligned}$$

3.3.2. Sufficient conditions for regularity

Definition 3 specifies the notion of regularity as the property that the local downward incentive constraints are the binding constraints in the principal's problem ($\mathbf{P}_\sigma^f(X, u^P, u^A)$). I now want to spell out sufficient conditions on the fundamentals of the model, i.e., the prior, the signal, and the payoff functions, such that regularity is satisfied. I start with the following restriction on the prior and the signal.

Definition 5. A signal σ induces decreasing inverse hazard rates for the prior f if $h_\sigma^f(\theta|s)$ is weakly decreasing in θ for all signal realizations $s \in S$.

The condition of decreasing hazard rates is a standard assumption in the literature on monopolistic screening and mechanism design. We can combine this condition with standard assumptions on the payoff functions to obtain the following result.[19]

Proposition 1. *Fix a prior f. Suppose the signals $\check{\sigma}$ and $\hat{\sigma}$ have full support and $\check{\sigma}$ induces decreasing inverse hazard rates for the prior f. The following two statements are equivalent*[20]*:*

1) $\Pi_{\check{\sigma}}^f(X, u^P, u^A) \geq \Pi_{\hat{\sigma}}^f(X, u^P, u^A)$ *for all* $(X, u^P, u^A) \in \mathcal{U}$ *such that* $X = [0, \bar{x}]$ *and*

[17] Define $[x]_+ \equiv \max\{x, 0\}$.

[18] Note that the expected value of this random variable is $\rho \cdot \frac{1-\rho}{\rho} + (1-\rho) \cdot \frac{\rho}{1-\rho} = 1$, and therefore equals the inverse hazard rate of the uniform prior.

[19] Cf. Section 7.3.2. in Fudenberg and Tirole (1991).

[20] A function $\psi : X \times Y \to \mathbb{R}$ with $X \subseteq \mathbb{R}$ and $Y \subseteq \mathbb{R}$ is (strictly) supermodular if $\psi(x, y) + \psi(x', y') \geq (>) \psi(x, y') + \psi(x', y)$ for all $x > x', y > y'$.

i) $u^P(\theta, x) = -c(x)$ *for all* $\theta \in \Theta$ *with* $c(0) = 0$,
ii) $u^A(\theta, x)$ *is strictly supermodular, increasing in* θ, *and* $u^A(\theta, 0) = 0$ *for all* θ,
iii) $-\Delta_v^A(\theta_i, x)$ *is supermodular.*

2) $\check{\sigma} \succeq^f_{HRS} \hat{\sigma}$.

I argue why condition 2) implies condition 1). Note first that some stochastic decision q maximizes the virtual surplus if and only if it induces only decisions x that maximize the deterministic virtual surplus function

$$\bar{u}(\theta_i, x) - h^f_\sigma(\theta_i|s_j)\Delta^A_u(\theta_i, x)$$

where $\Delta^A_u(\theta_i, x) \equiv u^A(\theta_{i+1}, x) - u^A(\theta_i, x)$ for $i = 1, \dots, m-1$ and $\Delta^A_u(\theta_m, \cdot) = 0$ in analogy to $\Delta^A_v(\theta_i, q)$. Secondly, observe that under conditions i), ii), and iii), the deterministic virtual surplus is strictly supermodular if the signal σ induces decreasing inverse hazard rates for the prior f. From these two observations, it follows that there exists a deterministic decision rule which maximizes the virtual surplus and is increasing in θ_i. As is well known in the literature on mechanism design, such a monotonic decision rule is also implementable – and therefore optimal – in the original problem ($\mathbf{P}^f_\sigma(X, u^P, u^A)$). Thus, any triple $(X, u^P, u^A) \in \mathcal{U}$ which satisfies conditions i), ii), and iii) is regular for the signal σ and the prior f. Theorem 2 then implies that condition 2) is sufficient for condition 1).

To see that condition 1) implies condition 2), recall the private, binary value, bilateral trade environment $(\tilde{X}, \tilde{u}^P, \tilde{u}^A)$ that I used to show necessity of the HRS order in the proof of Theorem 2. If we replace $\tilde{X}$ with $\tilde{X}' = [0, 1]$ and consider the obvious extensions of $\tilde{u}^P$ and $\tilde{u}^A$ to this domain, we obtain an equivalent environment which satisfies conditions i), ii), and iii) and ensures the necessity of the HRS order for the principal to prefer $\check{\sigma}$ over $\hat{\sigma}$.

3.3.3. Relation to other information orders

In this section, I relate the HRS order to other information orders proposed in the literature. I first provide a comparison to the information order to Blackwell (1951, 1953). Secondly, I describe the relation to the prior-independent information orders of Lehmann (1988), Kim (1995), and Kim (2023). Finally, I discuss the relationship to the prior-dependent information orders of Athey and Levin (2018) and Ganuza and Penalva (2010).

Relation to Blackwell's information order I obtain the following result regarding the relationship between the Blackwell order and the HRS order.

Proposition 2. *Given any two signals* $\check{\sigma}$ *and* $\hat{\sigma}$,

$$\check{\sigma} \succeq_B \hat{\sigma} \implies \check{\sigma} \succeq^f_{HRS} \hat{\sigma} \ \forall f \ \text{ and } \ \check{\sigma} \succeq^f_{HRS} \hat{\sigma} \ \forall f \ \not\Rightarrow \check{\sigma} \succeq_B \hat{\sigma}$$

The HRS order is weaker than Blackwell's order even if we do not impose any restrictions on the prior. If two signals are ranked by Blackwell's order, these signals are ranked in the same way by the HRS order for all priors. However, the reverse is not true. Even if two signals are ranked in the same way by the HRS order for all priors, they may not be ranked by Blackwell's order. The first part of the proposition is directly implied by Theorems 1 and 2. The second part is proved by means of an example.

Relation to Lehmann (1988), Kim (1995) and Kim (2023) Next, I turn to the relationship between the HRS order and the information orders of Lehmann (1988), Kim (1995) and Kim (2023). The latter three information orders are equivalent in the class of signals that satisfy the monotone likelihood ratio property.[21,22]

Definition 6. A signal σ satisfies the monotone likelihood ratio property if $\frac{g(s_j|\theta_k)}{g(s_j|\theta_i)}$ is increasing in j for any $k > i$.

I follow Jewitt (2007) in the definition of this information ordering.[23] For a given signal σ, consider for any two types $\theta_i, \theta_k \in \Theta$ the random variable which takes the value $g(s|\theta_k)/g(s|\theta_i)$ with probability $g(s|\theta_i)$. The cumulative distribution function of this random variable is given by

$$L_\sigma(z|\theta_i, \theta_k) \equiv \sum_{j=1}^{n} g(s_j|\theta_i)\mathbf{1}\left(\frac{g(s_j|\theta_k)}{g(s_j|\theta_i)} \le z\right).$$

We can then formulate the following information order.

[21] The equivalence follows from Propositions 1 and 10 in Jewitt (2007), or Proposition 3.1 in Dewatripont et al. (1999), and Theorem 1 in Kim (2023).
[22] Chi (2014) proposes a dispersion order which is also equivalent to the order of Lehmann (1988).
[23] Cf. Equation (2) in Jewitt (2007).

Definition 7. Consider two signals $\check{\sigma}$ and $\hat{\sigma}$ that satisfy the monotone likelihood ratio property. The signal $\check{\sigma}$ is ranked ahead of the signal $\hat{\sigma}$ in the information orders of Lehmann (1988), Kim (1995) and Kim (2023), denoted by $\check{\sigma} \succeq_{LK} \hat{\sigma}$, if for any $\theta_i, \theta_k \in \Theta$ with $k > i$ and all $z \geq 0$

$$\int_0^z L_{\check{\sigma}}(y|\theta_i, \theta_k) dy \geq \int_0^z L_{\hat{\sigma}}(y|\theta_i, \theta_k) dy.$$

Thus, two signals are ordered by the information orders of Lehmann (1988), Kim (1995) and Kim (2023) if one signal induces a more spread out distribution over likelihood ratios than the other signal.[24]

The following result establishes the relationship between the information order of Lehmann (1988) and the HRS order. Moreover, it relates the analysis of screening problems in Kim (2023) to the results in the current paper.[25]

Proposition 3. *Suppose the signals $\check{\sigma}$ and $\hat{\sigma}$ have full support and satisfy the monotone likelihood ratio property. The following statements are mutually equivalent:*

1) $\check{\sigma} \succeq_{LK} \hat{\sigma}$.
2) $\check{\sigma} \succeq^f_{HRS} \hat{\sigma} \ \forall f$.
3) $\underline{\Pi}^f_{\check{\sigma}}(X, u^P, u^A) \geq \underline{\Pi}^f_{\hat{\sigma}}(X, u^P, u^A)$ *for all* f *and* $(X, u^P, u^A) \in \mathcal{U}$ *such that* $X = [0, \bar{x}]$ *and*
   i) $u^P(\theta, x) = -c(x)$ *for all* $\theta \in \Theta$ *with* $c(0) = 0$*,*
   ii) $u^A(\theta, x)$ *is strictly supermodular, increasing in* θ*, and* $u^A(\theta, 0) = 0$ *for all* θ*.*

The proposition establishes the equivalence between the information orders of Lehmann (1988), Kim (1995), and Kim (2023) and the HRS order for all priors on the domain of signals that satisfy the monotone likelihood ratio property. If additional information about the prior is available, the HRS order can rank more signals, both within and outside of the domain of signals with the monotone likelihood ratio property.[26]

This equivalence can then be used to complement an important result in Kim (2023). In particular, Kim (2023) shows – for the domain of signals with the monotone likelihood ratio property – that the information order of Lehmann (1988) is a sufficient condition for the principal to prefer one signal over another for all screening problems in which the principal is restricted to offer deterministic mechanisms and conditions $3i)$ and $3ii)$ in Proposition 3 are satisfied. Using the equivalence of the Lehmann (1988) order with the HRS order for all priors, the necessity proof of Theorem 2 can be employed to show that the information order of Lehmann (1988) is not only a sufficient but also a necessary condition for the principal to prefer one signal over another in these screening problems.[27]

A comparison of Propositions 1 and 3 further clarifies how the results of this paper complement the result of Kim (2023) described in the previous paragraph. First, the two results pertain to different but overlapping domains of signals. While Kim (2023) focuses on prior-independent comparisons of signals with the monotone likelihood ratio property, Proposition 1 can make use of even partial knowledge about the prior and applies to signals which induce weakly decreasing inverse hazard rates. Second, the results complement each other by speaking to different sets of preferences and contracting options. Proposition 1 imposes condition iii) in addition to the conditions i) and ii) that are also required under Proposition 3. In turn, Proposition 3 restricts attention to deterministic mechanisms whereas there are no restrictions on the principal's choice of mechanisms in Proposition 1.[28]

Relation to Athey and Levin (2018) and Ganuza and Penalva (2010) Finally, I relate the HRS order to the monotone information order for decision makers with supermodular preferences (MIO-ND) in Athey and Levin (2018) and the integral precision order (IPO) of Ganuza and Penalva (2010). As the HRS order, both of these information orders are prior-dependent. MIO-ND applies to signals that induce posterior distributions which are ordered by first-order stochastic dominance. For this domain of signals, Athey and Levin (2018) show that any decision maker whose preferences are supermodular in a scalar action and a scalar state of the world prefers one signal over another signal if the two signals are ordered by MIO-ND.[29] Ganuza and Penalva (2010) define IPO as the order which ranks two signals if the distribution over posterior means induced by the first signal is a mean-preserving spread of the distribution induced by the second signal.[30]

---

24 Kim (1995) defines his information order using a local version of the likelihood ratio. Definition 7 is equivalent to the local formulation. If the random variable $\frac{g(\tilde{s}|\theta_{i+1})}{g(\tilde{s}|\theta_i)}$ is a mean-preserving spread of $\frac{\hat{g}(\tilde{s}|\theta_{i+1})}{\hat{g}(\tilde{s}|\theta_i)}$, then $\frac{g(\tilde{s}|\theta_{i+1})-g(\tilde{s}|\theta_i)}{g(\tilde{s}|\theta_i)} = \frac{\hat{g}(\tilde{s}|\theta_{i+1})}{\hat{g}(\tilde{s}|\theta_i)} - 1$ is a mean-preserving spread of $\frac{\hat{g}(\tilde{s}|\theta_{i+1})-\hat{g}(\tilde{s}|\theta_i)}{\hat{g}(\tilde{s}|\theta_i)} = \frac{\hat{g}(\tilde{s}|\theta_{i+1})}{\hat{g}(\tilde{s}|\theta_i)} - 1$.

25 Kim (2023) focuses on screening problems in which the principal offers deterministic mechanisms. Recall that $\underline{\Pi}^f_{\sigma}(X, u^P, u^A)$ denotes the principal's optimal payoff under this restriction.

26 For an example, see Lemma 9 in Appendix B.

27 I would like to thank an anonymous referee for pointing this out to me.

28 This relationship between preferences and contracting options is not surprising given that Strausz (2006) demonstrates that deterministic mechanisms are optimal in the set of general, i.e., potentially stochastic, mechanisms if there is no bunching. Condition iii) in Proposition 1 ensures the absence of bunching. Without this condition, Strausz (2006) shows that deterministic mechanisms may typically be suboptimal.

29 MIO-ND is formally defined in Section 3.3.1 in Athey and Levin (2018).

30 Cf. Definition 2 ii) in Ganuza and Penalva (2010).

Proposition 4. *The HRS order neither implies nor is implied by IPO or MIO-ND.*

I prove the proposition by two examples. In particular, I first specify two signals that are ranked by the criterion of MIO-ND but not by the HRS order for some prior. Thus, MIO-ND does not imply the HRS order. As MIO-ND implies the criterion of IPO,[31] the HRS order is also not implied by IPO. Second, I present two signals that can be ranked by the HRS order but not by the criterion of IPO under some prior. Thus, the HRS order does neither imply IPO nor MIO-ND.

4. Welfare analysis under HRS order

In this section, I demonstrate that the HRS order is a useful tool to study the welfare effects of changes in the information asymmetry between principal and agent. I provide sufficient conditions such that ex-ante welfare increases or decreases if the principal becomes better informed according to the HRS criterion. I then apply the sufficient conditions to provide new results regarding the welfare effects of nonlinear price discrimination.

4.1. Sufficient conditions for welfare comparison

First, I introduce a third party – a non-strategic player – to capture contractual externalities that may be relevant from a welfare perspective. For instance, in a vertically related industry consisting of a manufacturer and a retailer, consumers are a third party as their payoff depends on the final good price that results from the wholesale contract between manufacturer and retailer. In particular, let $u^T(\theta,x)$ denote the payoff of the third party for the decision $x \in X$ and the type $\theta \in \Theta$. I assume that $u^T(\theta,\cdot)$ is bounded and continuous in x for all θ. Thus, we have for any type θ a well-defined expected payoff $v^T(\theta,q) \equiv \int_X u^T(\theta,x)dq(x)$ for the stochastic decision $q \in \mathcal{Q}$. I denote the set of all combinations of such a function u^T and a triple $(X,u^P,u^A) \in \mathcal{U}$ by $\mathcal{U}'$.

I study welfare measures that are linear in the payoffs of principal, agent, and third party. Given the welfare weights $(\beta^P,\beta^A,\beta^T) \in \mathbb{R}^3$ on the payoffs of principal, agent and third party, a type θ and an allocation (q,t) generate a welfare of

$$w(\theta,q,t) \equiv \beta^P(v^P(\theta,q)+t) + \beta^A(v^A(\theta,q)-t) + \beta^T v^T(\theta,q).$$

If the principal's problem ($\mathbf{P}_\sigma^f(X,u^P,u^A)$) has multiple solutions, I assume that the principal chooses a welfare-optimal solution, i.e., the principal breaks ties in favor of welfare. Denote by $\mathcal{M}_\sigma^{*,f}(X,u^P,u^A)$ the set of solutions to ($\mathbf{P}_\sigma^f(X,u^P,u^A)$). The ex-ante welfare for the tuple (X,u^P,u^A,u^T) given the prior f and the signal $\sigma = (S,\{g(\cdot|\theta)\}_{\theta\in\Theta})$ is then

$$W_\sigma^f(X,u^P,u^A,u^T) \equiv \max_{\{(q(\cdot|s),t(\cdot|s))\}_{s\in S} \in \mathcal{M}_\sigma^{*,f}(X,u^P,u^A)} \sum_{i=1}^m \sum_{j=1}^n w(\theta_i, q(\theta_i|s_j), t(\theta_i|s_j)) g(s_j|\theta_i) f(\theta_i).$$

To provide a consistent comparison of the ex-ante welfare under different signals, I combine this assumption with the following adapted notion of regularity.

Definition 8. The quadruple $(X,u^P,u^A,u^T) \in \mathcal{U}'$ is regular for the prior f and the signal σ if there exists a welfare-optimal solution to ($\mathbf{R}_\sigma^f(X,u^P,u^A)$) that solves ($\mathbf{P}_\sigma^f(X,u^P,u^A)$). The set of all such quadruples is denoted by $\mathcal{U}_\sigma'^{R,f}$.

In the analysis of regular screening problems in Section 3.2, it was possible to allow for multiple solutions to the principal's problem as any of these solutions implied the same payoff to the principal. If our focus is on welfare measures that differ from the principal's payoff, two different solutions to the principal's problem might lead to different welfare outcomes. In this case, the signal could influence welfare beyond the signal's impact on the inverse hazard rate if – for a given type θ_i and a value of the inverse hazard rate z – the signal was to influence the selection of a virtual surplus-maximizing decision from the set

$$Q^*(\theta_i,z) \equiv \arg\max_{q\in\mathcal{Q}} \bar{v}(\theta_i,q) - z\Delta_v^A(\theta_i,q).$$

The assumption of welfare-optimal tie-breaking ensures that any selection from the set $Q^*(\theta_i,z)$ leads to the same ex-ante welfare. I then obtain the following result.

Lemma 2. *Suppose the signal σ has full support. Then, $(X,u^P,u^A,u^T) \in \mathcal{U}_\sigma'^{R,f}$ implies*

$$W_\sigma^f(X,v^P,v^A,v^T) = \sum_{i=1}^m \int_0^\infty \omega(\theta_i,z)dM_\sigma^f(z|\theta_i)f(\theta_i)$$

where $\omega(\theta_i,z) \equiv \beta^P\bar{v}(\theta_i,\rho(\theta_i,z)) - (\beta^P-\beta^A)z\Delta_v^A(\theta_i,\rho(\theta_i,z)) + \beta^T v^T(\theta_i,\rho(\theta_i,z))$ for some function $\rho : \Theta\times\mathbb{R} \to \mathcal{Q}$ with

[31] See Theorem 1(i) in Ganuza and Penalva (2010).

$$\rho(\theta_i, z) \in \arg \max_{q \in Q^*(\theta_i, z)} \beta^A z \Delta_v^A(\theta_i, q) + \beta^T v^T(\theta_i, q).$$

Under the adapted notion of regularity from Definition 8, Lemma 2 generalizes Lemma 1. The result builds on the well-known revenue equivalence condition according to which the agent's expected payoff under an optimal mechanism can be expressed as a function of the inverse hazard rate and the optimal decision. This allows me to write the ex-ante welfare as the expectation of a function $\omega(\theta_i, z)$ over the type and the signal realization. In particular, the function $\omega(\theta_i, z)$ is a linear combination of the joint surplus $\bar{v}(\theta_i, q)$, the agency cost $z\Delta_v^A(\theta_i, q)$, and the third party's payoff $v^T(\theta_i, q)$ – evaluated at an optimal decision $\rho(\theta_i, z)$.

As the formulation of ex-ante welfare in Lemma 2 suggests, the curvature of the function $\omega(\theta_i, \cdot)$ can be used to formulate sufficient conditions for ex-ante welfare to increase or decrease in the HRS order.

Theorem 3. *Given a prior f and two signals $\check{\sigma}$ and $\hat{\sigma}$ with full support, $\check{\sigma} \succeq_{HRS}^f \hat{\sigma}$ implies*

$$W_{\check{\sigma}}^f(X, v^P, v^A, v^T) \geq (\leq)\, W_{\hat{\sigma}}^f(X, v^P, v^A, v^T) \;\; \forall (X, v^P, v^A, v^T) \in \mathcal{U}'^{R,f}_{\check{\sigma}} \cap \mathcal{U}'^{R,f}_{\hat{\sigma}}$$

if $\omega(\theta_i, z)$ is convex (concave) in z for all $\theta_i \in \Theta$.

Theorem 3 provides simple sufficient conditions under which ex-ante welfare is increasing or decreasing as the principal's information improves along the HRS order. The result follows directly from the expression of ex-ante welfare in Lemma 2 and Jensen's inequality. If the function $\omega(\theta, z)$ is convex, ex-ante welfare increases under a more spread-out distribution of inverse hazard rates in analogue to the expected payoff of a risk-loving decision maker who prefers riskier lotteries. If $\omega(\theta, z)$ is concave, ex-ante welfare decreases along the HRS order as the expected payoff of a risk-averse decision maker would under a riskier lottery.

The curvature of the function $\omega(\theta, \cdot)$ captures the difference in the welfare effects for small and large distortions. The inverse hazard rate serves as a scaling factor on the wedge between joint surplus and virtual joint surplus due to asymmetric information. If the marginal loss in welfare due to a larger wedge is increasing as the wedge becomes larger, a more spread out distribution of the inverse hazard rate decreases welfare. This holds as smaller welfare losses from small values of the inverse hazard rate are more than offset by the larger welfare losses from large values of the inverse hazard rate. The reverse holds true if the marginal welfare loss decreases in the size of the wedge.

Before applying Theorem 3 to analyze the welfare effects of nonlinear price discrimination, I provide sufficient conditions for the tuple $(X, u^P, u^A, u^T) \in \mathcal{U}'$ to be regular given a prior f and a signal σ.

Proposition 5. *Fix a prior f. Suppose the signal σ has full support, induces decreasing inverse hazard rates for f and $(X, u^P, u^A, u^T) \in \mathcal{U}'$ with $X = [0, \bar{x}]$ and*

i) *$u^P(\theta, x) = -c(x)$ for all $\theta \in \Theta$ with $c(0) = 0$,*
ii) *$u^A(\theta, x)$ is strictly supermodular, increasing in θ, and $u^A(\theta, 0) = 0$ for all θ,*
iii) *$-\Delta_u^A(\theta_i, x)$ is supermodular.*

Then, $(X, u^P, u^A, u^T) \in \mathcal{U}'^{R,f}_{\sigma}$.

The proposition shows that the same conditions as those used in Proposition 1 can be invoked to ensure that the adapted notion of regularity from Definition 8 is satisfied.[32]

4.2. Application to discriminatory nonlinear pricing

A large part of the literature on the welfare effects of price discrimination assumes that the seller sets linear prices.[33] Herweg and Müller (2014) and Haghpanah and Siegel (2023) move beyond this restriction by studying a seller who uses information about the buyer's preferences when setting a nonlinear tariff. Haghpanah and Siegel (2023) consider a final good market in which the seller offers the buyer a menu of finitely many product alternatives. They show that it is generically possible to achieve a Pareto improvement of the market outcome by providing the seller with more information about the buyer. Herweg and Müller (2014) study the welfare effects of nonlinear price discrimination by an upstream firm on an intermediate good market. They compare the setting in which the upstream firm charges the downstream firm a nonlinear tariff based on the prior belief about the downstream firm's cost distribution with the situation in which the upstream firm observes an informative signal about the downstream firm's cost before offering a tariff. They show that the signal reduces expected total surplus if the final good market is covered.

In this section, I use the sufficient condition of Theorem 3 to derive novel results on the welfare effects of discriminatory nonlinear pricing in final and intermediate good markets and relate them to the contributions of Haghpanah and Siegel (2023) and Herweg and Müller (2014).

[32] Note that for an arbitrary prior f and signal σ, the set $\mathcal{U}'^{R,f}_{\sigma}$ is nonempty as it contains the binary value-bilateral trade payoff functions specified in the proof of Theorem 2.

[33] See Pigou (1920), Robinson (1933), Schmalensee (1981), Aguirre et al. (2010).

4.2.1. Nonlinear price discrimination on final good markets

I consider the following version of the classic model of Mussa and Rosen (1978). A seller can produce the quantity $x \in X = [0, \bar{x}]$ at a cost $c(x)$. A buyer's gross payoff from the quantity x is $\phi(x) + \theta x$ where $\theta \in \Theta \subseteq \mathbb{R} \sim f$ is the buyer's private information. Suppose the functions $c(\cdot)$ and $\phi(\cdot)$ are thrice differentiable with $\phi(0) = c(0) = 0$, the upper bound on quantity $\bar{x}$ is large enough in order not to be binding, and $\theta_{i+1} - \theta_i = \Delta\theta > 0$ for all $i = 1, \ldots, m-1$. The seller observes a signal σ before offering a direct mechanism to the buyer.[34],[35]

Proposition 6. *Consider the model of nonlinear pricing on final good markets of Mussa and Rosen (1978). Fix a prior f and two signals $\check{\sigma}$ and $\hat{\sigma}$ with full support, decreasing inverse hazard rates, and $\check{\sigma} \succeq^f_{HRS} \hat{\sigma}$. Total surplus is lower under the signal $\check{\sigma}$ if*

i) *$\phi(x) - c(x)$ is strictly concave,*
ii) *$\phi'(x) + \theta - c'(x)$ is log-concave for all θ, and*
iii) *production is strictly positive for all signal realizations and all types, i.e., $\phi'(0) + \min_S\{\theta_1 - 1/g^f_{\check{\sigma}}(\theta_1|s)\} > c'(0)$.*

The proposition provides conditions under which a better-informed seller reduces total surplus in the model of Mussa and Rosen (1978). Under these conditions, a ban on price discrimination increases total surplus, provided that the prior f has a decreasing inverse hazard rate. The conditions *i*) and *ii*) are relatively mild conditions on the functional forms of $\phi(\cdot)$ and $c(\cdot)$ which are satisfied by many standard specifications. Condition *iii*) is much more substantial. It requires that the seller never finds it optimal to exclude the lowest type θ_1 in order to save on information rents to the higher types.

I now use the example from Section 2.5 to show that condition *iii*) can be understood as a limit on the informativeness of the signal, and that – more generally – total surplus can be seen as U-shaped in the precision of the signal. First, I assume that $2\theta_1 > \theta_2$ so that there is strictly positive production for type θ_1 under the uniform prior. Production is strictly positive after both signal realizations s_1 and s_2 if $\theta_1 - \frac{\rho}{1-\rho}\Delta\theta > 0 \iff \rho < \bar{\rho} \equiv \frac{\theta_1}{\theta_2}$. If this condition holds, the type θ_2 receives the efficient quantity θ_2 for both signal realizations and the type θ_1 consumes a quantity of $\theta_1 - \frac{1-\rho}{\rho}\Delta\theta$ for the signal realization s_1 and $\theta_1 - \frac{\rho}{1-\rho}\Delta\theta$ for s_2. This yields an expected total surplus of

$$0.5\rho\left(\theta_1\left(\theta_1 - \tfrac{1-\rho}{\rho}\Delta\theta\right) - 0.5\left(\theta_1 - \tfrac{1-\rho}{\rho}\Delta\theta\right)^2\right)$$
$$+ 0.5(1-\rho)\left(\theta_1\left(\theta_1 - \tfrac{\rho}{1-\rho}\Delta\theta\right) - 0.5\left(\theta_1 - \tfrac{\rho}{1-\rho}\Delta\theta\right)^2\right) + 0.25\theta_2^2.$$

Note that the term $\theta_1(\theta_1 - z\Delta\theta) - 0.5(\theta_1 - z\Delta\theta)^2 = 0.5(\theta_1^2 - (z\Delta\theta)^2)$ is concave in z. Thus, a more spread out distribution of inverse hazard rates reduces total surplus by Jensen's inequality. In particular, the expected total surplus simplifies to

$$0.25(\theta_1^2 + \theta_2^2) - 0.25\Delta\theta^2\left(\frac{1 - 3\rho + 3\rho^2}{\rho(1-\rho)}\right),$$

and it is straightforward to check that the term $\frac{1-3\rho+3\rho^2}{\rho(1-\rho)}$ is increasing in ρ on the interval $[0.5, \bar{\rho}]$. Thus, total surplus is indeed decreasing in the precision of the signal as long as production is always positive. If we move beyond the realm of Proposition 6 and assume $\rho \geq \bar{\rho}$, the type θ_1 is excluded for the signal realization s_2. Thus, the total surplus in this case amounts to

$$0.5\rho\left(\theta_1\left(\theta_1 - \tfrac{1-\rho}{\rho}\Delta\theta\right) - 0.5\left(\theta_1 - \tfrac{1-\rho}{\rho}\Delta\theta\right)^2\right) + 0.25\theta_2^2$$
$$= 0.25(\theta_1^2 + \theta_2^2) - 0.25\left((1-\rho)\theta_1^2 + \frac{(1-\rho)^2}{\rho}\Delta\theta^2\right).$$

This term is clearly increasing in ρ. Thus, expected total surplus rises in the precision of the signal for $\rho \in [\bar{\rho}, 1]$ and attains the first-best surplus for the perfectly informative signal $\rho = 1$.

One can conclude that expected total surplus is U-shaped in the precision of the signal. For imprecise signals, there is no exclusion and an increase in the precision of the signal implies that the consumption of type θ_1 varies more strongly in the signal. As the surplus is a concave function in the quantity, this spread reduces total surplus. If the signal's precision increases beyond the threshold $\bar{\rho}$, the type θ_1 starts to be excluded after the signal realization s_2. A further increase in the precision of the signal then raises the expected total surplus as it reduces the distortion of the type θ_1 for the signal realization s_1 and makes exclusion of type θ_1 for the signal realization s_2 less likely. The U-shape of expected total surplus also implies that there exists some threshold $\tilde{\rho} \in (\bar{\rho}, 1)$ such that banning price discrimination – which is equivalent to an uninformative signal – leads to higher total surplus than allowing the monopolist to price discriminate based on the signal if, and only if, $\rho < \tilde{\rho}$. The critical threshold $\tilde{\rho}$ is given by

[34] A signal σ can be interpreted as a market segmentation where each signal realization corresponds to a market segment in which types are distribution according to the posterior $g^f_\sigma(\cdot|s)$. If each segment of a given market segmentation is further split up into market subsegments, the resulting finer market segmentation is equivalent to a signal σ' which Blackwell dominates the signal σ, and thus is also ranked ahead of σ in the HRS order.

[35] By the taxation principle, any direct deterministic mechanism is equivalent to a nonlinear tariff $T(x|s)$ (Hammond, 1979; Guesnerie, 1981).

$$\Delta\theta^2 = (1-\tilde{\rho})\theta_1^2 + \frac{(1-\tilde{\rho})^2}{\tilde{\rho}}\Delta\theta^2 \iff \tilde{\rho} = \frac{\sqrt{\left(\left(\frac{\theta_1}{\Delta\theta}\right)^2-1\right)^2+4}+\left(\frac{\theta_1}{\Delta\theta}\right)^2-3}{2\left(\left(\frac{\theta_1}{\Delta\theta}\right)^2-1\right)}.$$

It is noteworthy that as the difference between types becomes small, the critical threshold $\tilde{\rho}$ converges to one, i.e., only almost fully revealing signals may increase total surplus relative to a ban on price discrimination.[36]

Next, I discuss how Proposition 6 relates to Haghpanah and Siegel (2023) who show that welfare can generically be increased by providing carefully chosen additional information to the seller. The main difference between the environments of Proposition 6 and Haghpanah and Siegel (2023) is that they consider the case where the buyer receives one of finitely many product versions whereas Proposition 6 pertains to the case where the decision x is chosen from an interval.[37] Indeed, Haghpanah and Siegel (2023) exploit in their construction the fact that with finitely many production alternatives, a prior may be split into two posteriors such that the allocation under one of these posteriors remains the same as under the prior. In particular, if a posterior is very close to the prior, each type receives the same product version as under the prior. This is typically impossible to achieve if the production decision is selected from a continuum where any change in the principal's belief smoothly translates into a change of the induced allocation.[38] In Appendix C, I illustrate the difference between a finite and an infinite set of production possibilities using the example from Section 2.5.

4.2.2. Consumer surplus and wholesale price discrimination

I now study the following wholesale market model along the lines of Herweg and Müller (2014). A seller can produce a quantity $x \in X = [0, \bar{x}]$ at a linear cost of $(\kappa - \theta)x$ where $\theta \in \Theta \subseteq \mathbb{R} \sim f$ is private information of the seller and $\theta_{i+1} - \theta_i = \Delta\theta > 0$ for all $i = 1, \dots, m-1$. The seller can access a final market with strictly decreasing and thrice differentiable inverse demand function $P(x)$ via a platform. If a quantity x is sold on the final market, the seller generates the revenue $R(x) \equiv xP(x)$. I assume that the upper bound $\bar{x}$ is large enough in order not to be binding. The platform offers the seller a direct mechanism after having observed the realization of a signal σ about the seller's production costs.

Proposition 7. *Consider the model of nonlinear pricing on intermediate good markets. Fix a prior f and two signals $\breve{\sigma}$ and $\hat{\sigma}$ with full support, decreasing inverse hazard rates, and $\breve{\sigma} \succeq^f_{HRS} \hat{\sigma}$. Total surplus is lower under the signal $\breve{\sigma}$ if*

i) *revenue $R(x)$ is strictly concave,*
ii) *marginal revenue $R'(x)$ is concave, and*
iii) *final markets are covered for all signal realizations and types, i.e., $P'(0) > \kappa - \min_S\{\theta_1 - 1/g^f_{\breve{\sigma}}(\theta_1|s)\}$.*

The proposition captures the results of Herweg and Müller (2014) for covered final markets.[39] Given the prior has a decreasing inverse hazard rate, a concave marginal revenue function ensures that an uninformative signal – equivalent to a ban on wholesale price discrimination – leads to higher total surplus than any other signal with full support and decreasing inverse hazard rates. As in the case of nonlinear price discrimination on final good markets, the covered markets condition iii) plays a crucial role for the result. Instead of discussing the covered market conditions in more detail,[40] I complement the total surplus analysis perspective of Herweg and Müller (2014) by analyzing the effects of price discrimination on consumer surplus.

Proposition 8. *Consider the model of nonlinear pricing on intermediate good markets. Fix a prior f and two signals $\breve{\sigma}$ and $\hat{\sigma}$ with full support, decreasing inverse hazard rates, and $\breve{\sigma} \succeq^f_{HRS} \hat{\sigma}$. Consumer surplus is higher under the signal $\breve{\sigma}$ if*

i) *revenue $R(x)$ is strictly concave, and*
ii) *marginal demand $-P'(x)$ is not too log-convex, i.e., $log(-P'(x))'' \leq \frac{2}{x^2}$.*

The proposition shows that a better informed platform benefits the consumers. The conditions i) and ii) are quite mild. Concavity of the revenue function $R(x)$ is a standard assumption.[41] To see that condition ii) is also not overly restrictive, note that it is satisfied by any cubic demand curve $P(x) = 1 - ax - bx^2 - cx^3$ for $c \leq 0$, thereby including the cases of linear and quadratic demand, as well as the constant elasticity demand $P(x) = aq^{-b}$ for all cases where marginal revenue is positive, i.e., $b \leq 1$.

[36] For $\theta_1 = 2$ and $\theta_2 = 3$, we obtain $\tilde{\rho} = 0.77$, for $\theta_1 = 3$ and $\theta_2 = 4$, we have $\tilde{\rho} = 0.89$, and for $\theta_1 = 4$ and $\theta_2 = 5$, we get $\tilde{\rho} = 0.96$.

[37] Another difference consists in the fact that Proposition 6 imposes conditions on preferences and signals to ensure regularity and positive production for all types whereas Haghpanah and Siegel (2023) do not impose such conditions. However, in their Section V., Haghpanah and Siegel (2023) illustrate their main result using a model with linear preferences which closely corresponds to the model of nonlinear pricing considered in this subsection. There, they describe the construction of a Pareto-improving signal which generates positive and differentiated production decisions for all types. Thus, Haghpanah and Siegel (2023) do not require exclusion or bunching of types to obtain their result.

[38] Haghpanah and Siegel (2023) discuss this point in their appendix A.4.

[39] See their Proposition 8 which focuses on the case of linear demand functions.

[40] Most of the comments made in the previous section on price discrimination in final good markets still apply.

[41] This condition may be further weakened to strict quasiconcavity of $R(x) - (\kappa - \theta + z\Delta\theta_i)x$.

The proposition uncovers a tension between the welfare standards of total surplus and consumer surplus. Joining the conditions of Propositions 7 and 8, a competition authority's evaluation regarding the desirability of price discrimination by a platform hinges on the welfare standard it employs. In particular, the authority's optimal policy depends on the weight it attaches to those firms in the vertical industry that have little bargaining power. Their loss from a more informative signal exceeds the consumers' gain.

5. Conclusion

This paper studies robust comparisons of signals in screening problems with respect to the principal's payoff as well as other welfare measures.

I first study information orders from the perspective of the principal. As a benchmark result, I show that the principal prefers one signal over another for any possible combination of preferences of principal and agent if and only if the signals can be ordered by the information order of Blackwell (1951, 1953). I then restrict attention to the case of regular preferences in which the principal's problem can be solved using the Myersonian approach. I propose the Hazard Rate Spread (HRS) order according to which one signal is ranked ahead of another signal if – for any type – the distribution over inverse hazard rates induced by the first signal is a mean-preserving spread of the distribution induced by the latter signal. I show that the principal prefers one signal over another for any possible regular combination of preferences if and only if the signals are ranked according to the HRS order.

Moreover, I demonstrate that the HRS order is a useful tool to compare signals in screening problems with respect to other welfare measures than the principal's payoff. I consider the effect of a better informed principal in the sense of the HRS order on welfare measures that are linear in the payoffs of the principal, agent, and potentially affected third parties. I provide sufficient conditions such that ex-ante welfare is increasing or decreasing in the HRS order. I apply these sufficient conditions to study the welfare effect of price discrimination with nonlinear tariffs in wholesale and final good markets.

The current paper focuses on screening models with a single agent and static private information. While central elements of this paper may carry over to environments with many agents or dynamic private information, new aspects arise. If a principal interacts with several agents, a signal affects what agents know about each other. If a principal interacts with an agent whose type changes over time, a signal may influence the principal's knowledge of the stochastic process governing the agent's type, and thereby affect the ability to extract dynamic information rents. The analysis of these interesting cases is left for future research.

Declaration of competing interest

The author declares that he has no relevant or material financial interests that relate to the research described in this paper.

Appendix A. Proofs

A.1. Proof of Theorem 1

Lemma 3. $\check{\sigma} \succeq_B \hat{\sigma} \implies \Pi^f_{\check{\sigma}}(X, u^P, u^A) \geq \Pi^f_{\hat{\sigma}}(X, u^P, u^A)\ \forall (X, u^P, u^A) \in \mathcal{U}$.

Proof. The following proof is standard. I provide it for the sake of completeness. Suppose $\check{\sigma} \succeq_B \hat{\sigma}$. Take some arbitrary $(X, u^P, u^A) \in \mathcal{U}$. I show that any ex-ante payoff attainable under $\hat{\sigma}$ can also be achieved under $\check{\sigma}$. There exists a matrix B with the properties of Definition 1. Let $\{(\hat{q}(\cdot|\hat{s}_k), \hat{t}(\cdot|\hat{s}_k))\}_{k=1,\dots,\hat{n}}$ be a collection of feasible mechanisms under $\hat{\sigma}$. We construct a collection of mechanisms $\{(\check{q}(\cdot|s_j), \check{t}(\cdot|s_j))\}_{j=1,\dots,\check{n}}$ under $\check{\sigma}$ as follows. For each realization $\check{s}_j \in \check{S}$, $\check{q}(\cdot|s_j)$ is obtained from randomizing over the set $\{\hat{q}(\cdot|\hat{s}_k)\}_{k=1,\dots,\hat{n}}$ according to the probability distribution $\{b_{j1}, \dots, b_{j\hat{n}}\}$ and $\check{t}(\cdot|\check{s}_j) \equiv \sum_{k=1}^{\hat{n}} b_{jk}\hat{t}(\cdot|\hat{s}_k)$. This mechanism is in $\mathcal{M}_{\check{\sigma}}(X, u^A)$ by the fact that $(\hat{q}(\cdot|\hat{s}_k), \hat{t}(\cdot|\hat{s}_k))$ is in $\mathcal{M}_{\hat{\sigma}}(X, u^A)$. Moreover,

$$
\begin{aligned}
&\sum_{i=1}^{m}\sum_{j=1}^{\check{n}} \left(v^P(\theta_i, \check{q}(\theta_i|s_j)) + \check{t}(\theta_i|s_j)\right) \check{g}(\check{s}_j|\theta_i) f(\theta_i) \\
&\quad = \sum_{i=1}^{m}\sum_{j=1}^{\check{n}}\sum_{k=1}^{\hat{n}} b_{jk} \left(v^P(\theta_i, \hat{q}(\theta_i|\hat{s}_k)) + \hat{t}(\theta_i|\hat{s}_k)\right) \check{g}(\check{s}_j|\theta_i) f(\theta_i) \\
&\quad = \sum_{i=1}^{m}\sum_{k=1}^{\hat{n}} \left(v^P(\theta_i, \hat{q}(\theta_i|\hat{s}_k)) + \hat{t}(\theta_i|\hat{s}_k)\right) \left(\sum_{j=1}^{\check{n}} b_{jk}\check{g}(\check{s}_j|\theta_i)\right) f(\theta_i) \\
&\quad = \sum_{i=1}^{m}\sum_{k=1}^{\hat{n}} \left(v^P(\theta_i, \hat{q}(\theta_i|\hat{s}_k)) + \hat{t}(\theta_i|\hat{s}_k)\right) \hat{g}(\theta_i, \hat{s}_k).
\end{aligned}
$$

Thus, $\Pi^f_{\check{\sigma}}(X, u^P, u^A) \geq \Pi^f_{\hat{\sigma}}(X, u^P, u^A)$. □

To prove necessity of Blackwell's ordering, consider a decision maker whose payoff $u^D(\theta, x)$ depends on the decision $x \in X \subseteq \mathbb{R}^k$ and the state $\theta \in \Theta$ such that $u^D(\theta, x)$ is continuous in x and bounded for all $\theta \in \Theta$. Let $v^D(\theta, q) = \int_X dm(\theta, x) dq(x)$ denote the expected gross payoff from a stochastic decision $q \in Q = \Delta X$. To simplify the exposition, normalize the decision maker's payoff from the outside option $O \in X$ to zero. Let $\mathcal{U}^D$ denote the set of all pairs (X, u^D) that satisfy these conditions. Before making a choice, the decision maker observes the realization of a signal σ. Let $\{q^D(s)\}_{s\in S}$ be a collection of decisions for the signal σ and let $\mathcal{D}_\sigma$ be the set of all such collections. The decision maker's optimal ex-ante payoff from a signal σ given some $(X, u^D) \in \mathcal{U}^D$ and a prior f is given by

$$\Pi^{D,f}_\sigma(X, u^D) \equiv \max_{\mathcal{D}_\sigma} \sum_{i=1}^m \sum_{j=1}^n v^D(\theta_i, q^D(s_j)) g(s_j|\theta_i) f(\theta_i).$$

Blackwell's theorem can then be stated as follows.

Blackwell's Theorem. *Given a prior f and two signals $\check{\sigma}$ and $\hat{\sigma}$, we have*

$$\Pi^{D,f}_{\check{\sigma}}(X, u^D) \geq \Pi^{D,f}_{\hat{\sigma}}(X, u^D) \ \ \forall (X, u^D) \in \mathcal{U}^D \iff \check{\sigma} \succeq_B \hat{\sigma}.$$

Proof. Crémer (1982) provides a simple proof of this result. □

I start the proof of necessity with the following result.

Lemma 4. *Given f, σ, (X, u^D), and $\alpha > 0$, there exists a constant $I \geq 0$ such that*

$$\Pi^{D,f}_\sigma(X, u^D) - \alpha I \leq \Pi^f_\sigma(X, (1+\alpha)u^D, -\alpha u^D) \leq \Pi^{D,f}_\sigma(X, u^D).$$

Proof. Let $v^P(\theta, q) = (1+\alpha) v^D(\theta, q)$ and $v^A(\theta, q) = -\alpha v^D(\theta, q)$ for $\alpha > 0$. The ex-ante payoff for any collection of feasible mechanisms $\{(q(\cdot|s_j), t(\cdot|s_j))\}_{j=1,\dots,n}$ satisfies

$$\sum_{i=1}^m \sum_{j=1}^n \left(v^P(\theta_i, q(\theta_i|s_j)) + t(\theta_i|s_j) \right) g^f(\theta_i, s_j) \leq \sum_{i=1}^m \sum_{j=1}^n v^D(\theta_i, q(\theta_i|s_j)) g^f(\theta_i, s_j)$$

as $v^P(\theta, q(\theta|s)) = \bar{v}(\theta, q(\theta|s)) - v^A(\theta, q(\theta|s))$, $\bar{v}(\theta, q) = v^D(\theta, q)$, and $v^A(\theta, q(\theta|s)) \geq t(\theta|s))$ by ($PC_{\sigma:s}(X, u^A)$). ($IC_{\sigma:s}(X, u^A)$) implies for any pair $\theta, \theta' \in \Theta_\sigma(s)$

$$v^A(\theta, q(\theta|s)) - v^A(\theta', q(\theta|s)) \geq v^A(\theta, q(\theta'|s)) - v^A(\theta', q(\theta'|s)).$$

Due to $v^A(\theta, q) = -\alpha v^D(\theta, q)$, this is equivalent to

$$v^D(\theta, q(\theta|s)) + v^D(\theta', q(\theta'|s)) \leq v^D(\theta', q(\theta|s)) + v^D(\theta, q(\theta'|s)).$$

Thus, $v^D(\theta, q(\theta|s)) \leq v^D(\theta, q(\theta'|s))$ or $v^D(\theta', q(\theta'|s)) \leq v^D(\theta', q(\theta'|s))$. It follows that for any pair of types $\theta, \theta' \in \Theta_\sigma(s)$, we can replace $q(\theta|s)$ with $q(\theta'|s)$ – or vice versa – and increase $v^D(\theta, q)$. Hence, there exists a collection of decisions $\{q^D(s)\}_{s\in S}$ such that

$$\sum_{i=1}^m \sum_{j=1}^n v^D(\theta_i, q(\theta_i|s_j)) g^f_\sigma(\theta_i, s_j) \leq \sum_{i=1}^m \sum_{j=1}^n v^D(\theta_i, q^D(s_j)) g^f_\sigma(\theta_i, s_j) \leq \Pi^{D,f}_\sigma(X, u^D).$$

Thus, $\Pi^f_\sigma(X, (1+\alpha)u^D, -\alpha u^D) \leq \Pi^{D,f}_\sigma(X, u^D)$.

As a next step, I prove that $\Pi^f_\sigma(X, (1+\alpha)u^D, -\alpha u^D) \geq \Pi^{D,f}_\sigma(X, u^D) - \alpha I$ where

$$I = I^f_\sigma \equiv \sup_{\mathcal{D}_\sigma} \sum_{i=1}^m \sum_{j=1}^n \left(\max_{\theta\in\Theta} v^D(\theta, q^D(s_j)) - v^D(\theta_i, q^D(s_j)) \right) g^f_\sigma(\theta_i, s_j).$$

Note that any collection of decisions $\{q^D(s)\}_{s\in S}$ forms a collection of feasible mechanism together with the constant transfers $t(\theta|s) = \min_{\theta\in\Theta} v^A(\theta, q^d(s))$. We therefore have for any $\{q^D(s)\}_{s\in S} \in \mathcal{D}_\sigma$

$$\begin{aligned}
&\sum_{i=1}^m \sum_{j=1}^n v^D(\theta_i, q^D(s_j)) g^f_\sigma(\theta_i, s_j) - \alpha I^f_\sigma \\
&\quad \leq \sum_{i=1}^m \sum_{j=1}^n \left(v^D(\theta_i, q^D(s_j)) - \alpha \left(\max_{\theta\in\Theta} v^D(\theta, q^D(s_j)) - v^D(\theta_i, q^D(s_j)) \right) \right) g^f_\sigma(\theta_i, s_j) \\
&\quad = \sum_{i=1}^m \sum_{j=1}^n \left(\bar{v}(\theta_i, q^D(s_j)) - \left(v^A(\theta_i, q^D(s_j)) - \min_{\theta\in\Theta} v^A(\theta, q^D(s_j)) \right) \right) g^f_\sigma(\theta_i, s_j) \\
&\quad \leq \Pi^f_\sigma(X, (1+\alpha)u^D, -\alpha u^D).
\end{aligned}$$

Thus, $\Pi_\sigma^f(X,(1+\alpha)u^D,-\alpha u^D) \geq \Pi_\sigma^{D,f}(X,u^D) - \alpha I$. $\square$

Necessity of Blackwell's ordering then follows from this result.

Lemma 5. *For a prior f and two signals $\check{\sigma}$ and $\hat{\sigma}$, we have*

$$\Pi_{\check{\sigma}}^f(X,u^P,u^A) \geq \Pi_{\hat{\sigma}}^f(X,u^P,u^A) \;\; \forall (X,u^P,u^A) \in \mathcal{U}$$
$$\implies \Pi_{\check{\sigma}}^{D,f}(X,u^D) \geq \Pi_{\hat{\sigma}}^{D,f}(X,u^D) \;\; \forall (X,u^D) \in \mathcal{U}^D.$$

Proof. I prove the following contrapositive of this statement.

$$\exists (X,v^D) \in \mathcal{U}^D \text{ s.t. } \Pi_{\hat{\sigma}}^{D,f}(X,u^D) > \Pi_{\check{\sigma}}^{D,f}(X,u^D)$$
$$\implies \exists (X,u^P,u^A) \in \mathcal{U} \text{ s.t. } \Pi_{\hat{\sigma}}^f(X,u^P,u^A) > \Pi_{\check{\sigma}}^f(X,u^P,u^A).$$

Suppose (X,u^D) satisfies $\Pi_{\check{\sigma}}^{D,f}(X,u^D) < \Pi_{\hat{\sigma}}^{D,f}(X,u^D)$. Let $u^P(\theta,x) = (1+\alpha)u^D(\theta,x)$ and $u^A(\theta,x) = -\alpha u^D(\theta,x)$ for $\alpha > 0$. By Lemma 4, we obtain for $\alpha < (\Pi_{\hat{\sigma}}^{D,f}(X,u^D) - \Pi_{\check{\sigma}}^{D,f}(X,u^D))/I_{\hat{\sigma}}^f$

$$\Pi_{\hat{\sigma}}^f(X,(1+\alpha)u^D,-\alpha u^D) \geq \Pi_{\hat{\sigma}}^{D,f}(X,u^D) - \alpha I_{\hat{\sigma}}^f > \Pi_{\check{\sigma}}^{D,f}(X,u^D) \geq \Pi_{\check{\sigma}}^f(X,(1+\alpha)u^D,-\alpha u^D). \quad \square$$

A.2. Proof of Lemma 1

Fix a prior f and a signal σ with full support. Let $Z(\theta) \equiv \{z \geq 0 : \exists s \in S \text{ s.t. } h_\sigma^f(\theta|s) = z\} = \{z_1(\theta), \dots, z_{o(\theta)}(\theta)\}$ with $z_\ell(\theta) < z_{\ell+1}(\theta)$ for $\ell \in \{1,\dots,o(\theta)-1\}$. Let $z_0 < z_1(\theta)$. Define

$$m_\sigma^f(z_\ell(\theta)|\theta) \equiv \sum_{j=1}^n g(s_j|\theta)\mathbf{1}(h(\theta|s_j) = z_\ell(\theta)).$$

For any (X,u^P,u^A), we have

$$\begin{aligned}
\Pi_\sigma^f(X,u^P,u^A) &= \sum_{j=1}^n g_\sigma^f(s_j)\sum_{i=1}^m g_\sigma^f(\theta_i|s_j)\left(\max_Q \bar{v}(\theta_i,q) - h(\theta_i|s_j)\Delta_v^A(\theta_i,q)\right)\\
&= \sum_{i=1}^m f(\theta_i)\sum_{\ell=1}^{o(\theta_i)} m_\sigma^f(z_\ell(\theta_i)|\theta_i)\left(\max_Q \bar{v}(\theta_i,q) - z_\ell(\theta_i)\Delta_v^A(\theta_i,q)\right)\\
&= \sum_{i=1}^m f(\theta_i)\sum_{\ell=1}^{o(\theta_i)} (M_\sigma^f(z_\ell(\theta_i)|\theta_i) - M_\sigma^f(z_{\ell-1}(\theta_i)|\theta_i))\left(\max_Q \bar{v}(\theta_i,q) - z_\ell(\theta_i)\Delta_v^A(\theta_i,q)\right)\\
&= \sum_{i=1}^m f(\theta_i)\int_0^\infty \left(\max_Q \bar{v}(\theta_i,q) - z\Delta_v^A(\theta_i,q)\right) dM_\sigma^f(z|\theta_i). \quad \square
\end{aligned}$$

A.3. Proof of Theorem 2

I first prove that $\check{\sigma} \succeq_{HRS}^f \hat{\sigma} \implies \Pi_{\check{\sigma}}^f(X,u^P,u^A) \geq \Pi_{\hat{\sigma}}^f(X,u^P,u^A) \,\forall (X,u^P,u^A) \in \mathcal{U}_\sigma^{R,f}$.

I show that $\omega^P(\theta_i,z) \equiv \max_Q \bar{v}(\theta_i,q) - z\Delta_v^A(\theta_i,q)$ is convex in z for any $\theta_i \in \Theta$. Fix a type $\theta_i \in \Theta$. Consider $z', z'' \geq 0$ and define $z''' \equiv \lambda z' + (1-\lambda)z''$ for $\lambda \in [0,1]$. Let q', q'', and q''' be optimal decisions for z', z'', and z''' respectively. Convexity follows from

$$\begin{aligned}
\omega^P(\theta_i,z''') &= \bar{v}(\theta_i,q''') - z'''\Delta_v^A(\theta_i,q''')\\
&= \lambda(\bar{v}(\theta_i,q''') - z'\Delta_v^A(\theta_i,q''')) + (1-\lambda)(\bar{v}(\theta_i,q''') - z''\Delta_v^A(\theta_i,q'''))\\
&\leq \lambda(\bar{v}(\theta_i,q') - z'\Delta_v^A(\theta_i,q')) + (1-\lambda)(\bar{v}(\theta_i,q'') - z''\Delta_v^A(\theta_i,q''))\\
&= \lambda\omega^P(\theta_i,z') + (1-\lambda)\omega^P(\theta_i,z'').
\end{aligned}$$

For $\check{\sigma} \succeq_{HRS}^f \hat{\sigma}$, $M_{\check{\sigma}}^f(z|\theta)$ is a mean-preserving spread of $M_{\hat{\sigma}}^f(z|\theta)$ and it follows directly from Theorem 3.A.1 in Shaked and Shanthikumar (2007) that for each $\theta \in \Theta$

$$\int \omega^P(\theta,z)dM_{\check{\sigma}}^f(z|\theta) \geq \int \omega^P(\theta,z)dM_{\hat{\sigma}}^f(z|\theta).$$

For $(X,u^P,u^A) \in \mathcal{U}_{\check{\sigma}}^{R,f}$, this implies together with Lemma 1 that

$$\Pi^f_{\hat{\sigma}}(X,u^P,u^A) \leq \sum_{i=1}^m \int \omega^P(\theta_i,z) dM^f_{\hat{\sigma}}(z|\theta_i) f(\theta_i) \leq \sum_{i=1}^m \int \omega^P(\theta_i,z) dM^f_{\check{\sigma}}(z|\theta_i) f(\theta_i) = \Pi^f_{\check{\sigma}}(X,u^P,u^A).$$

It remains to show that $\Pi^f_{\check{\sigma}}(X,u^P,u^A) \geq \Pi^f_{\hat{\sigma}}(X,u^P,u^A)\ \forall (X,u^P,u^A) \in \mathcal{U}^{R,f}_\sigma \implies \check{\sigma} \succeq^f_{HRS} \hat{\sigma}$. I prove the contrapositive statement

$$\check{\sigma} \not\succeq^f_{HRS} \hat{\sigma} \implies \exists (X,u^P,u^A) \in \mathcal{U}^{R,f}_\sigma\ \Pi^f_{\check{\sigma}}(X,u^P,u^A) < \Pi^f_{\hat{\sigma}}(X,u^P,u^A).$$

Suppose $\check{\sigma} \not\succeq_{HRS} \hat{\sigma}$. Thus, there exist $\tilde{\theta} \in \Theta$ and $\tilde{z} > 0$ such that

$$\int_0^{\tilde{z}} M^f_{\check{\sigma}}(z|\tilde{\theta}) dz < \int_0^{\tilde{z}} M^f_{\hat{\sigma}}(z|\tilde{\theta}) dz.$$

Let $\tilde{X} = \{0,1\}$, $\tilde{u}^P(\theta_i,x) = 0$, and $\tilde{u}^A(\theta_i,x) = v(\theta_i)x$ with

$$v(\theta_i) = \begin{cases} 1 & \text{if} \quad \theta_i > \tilde{\theta}, \\ \hat{z}/(1+\hat{z}) & \text{if} \quad \theta_i = \tilde{\theta}, \\ -1 & \text{if} \quad \theta_i < \tilde{\theta}. \end{cases}$$

Fix a signal $\sigma \in \{\check{\sigma},\hat{\sigma}\}$ and a signal realization $s \in S$. For $\theta_i < \tilde{\theta}$, the virtual surplus satisfies $v(\theta_i)q - z(v(\theta_{i+1}) - v(\theta_i))q \leq -q$. Thus, any solution to the relaxed problem features $q(\theta_i|s) = 0$ for $\theta_i < \tilde{\theta}$. For $\theta_i > \tilde{\theta}$, the virtual surplus satisfies $v(\theta_i)q - z(v(\theta_{i+1}) - v(\theta_i))q = q$. Hence, $q(\theta_i|s) = 1$ for $\theta_i > \tilde{\theta}$ in any solution to the relaxed problem. For $\theta_i = \tilde{\theta}$, the virtual surplus simplifies to $\frac{\hat{z}-z}{1+\hat{z}}q$. Thus, $q(\tilde{\theta}|s) = \mathbf{1}(h^f_\sigma(\tilde{\theta}|s) \leq \tilde{z})$ in the solution to the relaxed problem. Note that any solution to the relaxed problem is nondecreasing in $v(\theta_i)$. It is therefore standard to show that the solution to the relaxed problem coincides with the solution to the original problem. For the calculation of the principal's ex-ante payoff, note that $\int_0^\infty \omega^P(\theta,z) d\tilde{M}^f_{\tilde{\sigma}}(z|\theta) = 0$ for any $\theta < \tilde{\theta}$. For any $\theta > \tilde{\theta}$, we have $\int_0^\infty \omega^P(\theta,z) dM^f_{\tilde{\sigma}}(z|\theta) = 1$. Thus, the ex-ante payoffs with σ and $\hat{\sigma}$ may only differ for $\theta = \tilde{\theta}$ where

$$\int_0^\infty \omega^P(\tilde{\theta},z) dM^f_{\tilde{\sigma}}(z|\tilde{\theta}) = \int_0^{\tilde{z}} \frac{\tilde{z}-z}{1+\tilde{z}} dM^f_{\tilde{\sigma}}(z|\tilde{\theta}) = \frac{\int_0^{\tilde{z}} M^f_{\tilde{\sigma}}(z|\tilde{\theta}) dz}{1+\tilde{z}}.$$

Hence, $\Pi^f_{\hat{\sigma}}(\tilde{X},\tilde{u}^P,\tilde{u}^A) > \Pi^f_{\check{\sigma}}(\tilde{X},\tilde{u}^P,\tilde{u}^A)$ follows from $\int_0^{\tilde{z}} M^f_{\check{\sigma}}(z|\tilde{\theta}) dz < \int_0^{\tilde{z}} M^f_{\hat{\sigma}}(z|\tilde{\theta}) dz$. □

A.4. Proof of Proposition 1

To prove that 2) implies 1), I show that any $(X,u^P,u^A) \in \mathcal{U}$ that satisfies conditions i), ii) and iii) lies in $\mathcal{U}^{R,f}_\sigma$ if σ induces decreasing inverse hazard rates. Note first that the solutions to the relaxed problem ($\mathbf{R}^f_\sigma(X,u^P,u^A)$) are those collections of mechanisms which consist of decision rules that induce for each s_j and θ_i decisions in the set

$$Q^{*,f}_\sigma(\theta_i|s_j) \equiv \arg\max_{q \in Q} \bar{v}(\theta_i,q) - h^f_\sigma(\theta_i|s_j) \Delta^A_v(\theta_i,q)$$

and transfer rules that are induced by the (binding) constraints of ($\mathbf{R}^f_\sigma(X,u^P,u^A)$). Due to the linearity of the virtual surplus in q, the set $Q^{*,f}_\sigma(\theta_i|s_j)$ consists of all q whose support is included in the set

$$X^{*,f}_\sigma(\theta_i|s_j) \equiv \arg\max_{x \in X} \bar{u}(\theta_i,x) - h^f_\sigma(\theta_i|s_j) \Delta^A_u(\theta_i,x)$$

where $\Delta^A_u(\theta_i,x) \equiv u^A(\theta_{i+1},x) - u^A(\theta_i,x)$ for $i = 1,\ldots,m-1$ and $\Delta^A_u(\theta_m,\cdot) = 0$. Consider the collection of mechanisms which consists for all $s_j \in S$ of the deterministic decision rules $q^*(\cdot|s_j) = \delta_{x^*(\cdot|s_j)}$ where $x^*(\theta_i|s_j) \in X^{*,f}_\sigma(\theta_i|s_j)$ for all θ_i and the transfer rules $t^*(\cdot|s_j)$ induced by the (binding) constraints of ($\mathbf{R}^f_\sigma(X,u^P,u^A)$). Note that this collection solves the relaxed problem ($\mathbf{R}^f_\sigma(X,u^P,u^A)$). If the inverse hazard rate $h^f_\sigma(\theta_i|s)$ is decreasing in θ_i, conditions i), ii), and iii) from Proposition 1 ensure that the deterministic virtual surplus function $u^A(\theta_i,x) - c(x) - h^f_\sigma(\theta_i|s)\Delta^A_u(\theta_i,x)$ is strictly supermodular. Due to Topkis' monotonicity theorem, any selection from the set $X^{*,f}_\sigma(\theta_i|s_j)$ is weakly increasing in θ_i.[42] Standard results from the literature then imply that the collection of deterministic mechanisms specified above also satisfies all constraints in the original problem ($\mathbf{P}^f_\sigma(X,u^P,u^A)$).[43] It is therefore also a solution to ($\mathbf{P}^f_\sigma(X,u^P,u^A)$). Thus, any $(X,u^P,u^A) \in \mathcal{U}$ that satisfies conditions i), ii) and iii) lies in $\mathcal{U}^{R,f}_\sigma$ if σ induces decreasing inverse hazard rates. It follows from Theorem 2 that 2) implies 1).

[42] Cf. Theorem 1 in Amir (2005).

[43] Cf. Theorem 7.3 in Fudenberg and Tirole (1991).

To prove that 1) implies 2), we can follow exactly the same steps as in the proof of necessity of the HRS order in Theorem 2 with the slight change that we use the triple $(\tilde{X}', \tilde{u}'^P, \tilde{u}'^A) \in \mathcal{U}$ instead of the triple $(\tilde{X}, \tilde{u}^P, \tilde{u}^A) \in \mathcal{U}$ and we set $\tilde{X}' = \Delta 0, 1 = [0,1]$, $\tilde{u}'^i = \tilde{v}^i$ for $i \in \{P, A\}$ such that $(\tilde{X}', \tilde{u}'^P, \tilde{u}'^A)$ satisfies conditions i), ii), and iii) in the proposition. □

A.5. Proof of Proposition 2

Theorems 1 and 2 imply for any prior f

$$\begin{aligned}
\breve{\sigma} \succeq_B \hat{\sigma} &\implies \Pi^f_{\breve{\sigma}}(X, u^P, u^A) \geq \Pi^f_{\hat{\sigma}}(X, u^P, u^A) \;\; \forall (X, u^P, u^A) \in \mathcal{U} \\
&\implies \Pi^f_{\breve{\sigma}}(X, u^P, u^A) \geq \Pi^f_{\hat{\sigma}}(X, u^P, u^A) \;\; \forall (X, u^P, u^A) \in \mathcal{U}^{R,f}_{\breve{\sigma}} \\
&\implies \breve{\sigma} \succeq^f_{HRS} \hat{\sigma}.
\end{aligned}$$

Lemmas 6 and 9 in Appendix B prove by example that the reverse statement is not true. □

A.6. Proof of Proposition 3

The proof consists of two steps. I first prove that condition 2) implies condition 1). In a second step, I prove that condition 3) implies condition 2). The result then follows as Proposition 7 in Kim (2023) shows that condition 1) implies condition 3).

As a first step, I prove that for any two signals $\breve{\sigma}$ and $\hat{\sigma}$ that satisfy the monotone likelihood ratio property (MLRP), $\breve{\sigma} \succeq^f_{HRS} \hat{\sigma}$ for all f implies $\breve{\sigma} \succeq_{LK} \hat{\sigma}$. For any signal σ that satisfies MLRP, we have

$$L_\sigma(z|\theta_i, \theta_k) = \begin{cases} 0 & \text{if } z < \frac{g(s_1|\theta_k)}{g(s_1|\theta_i)}, \\ G(s_j|\theta_i) & \text{if } z \in \left[\frac{g(s_j|\theta_k)}{g(s_j|\theta_i)}, \frac{g(s_{j+1}|\theta_k)}{g(s_{j+1}|\theta_i)}\right) \text{ and } 1 \leq j < n, \\ 1 & \text{if } z \geq \frac{g(s_n|\theta_k)}{g(s_n|\theta_i)}. \end{cases}$$

Thus,

$$\int_0^z L_\sigma(y|\theta_i, \theta_k) dy = \begin{cases} 0 & \text{if } z < \frac{g(s_1|\theta_k)}{g(s_1|\theta_i)}, \\ G(s_j|\theta_i) z - G(s_j|\theta_k) & \text{if } z \in \left[\frac{g(s_j|\theta_k)}{g(s_j|\theta_i)}, \frac{g(s_{j+1}|\theta_k)}{g(s_{j+1}|\theta_i)}\right) \text{ and } 1 \leq j < n, \\ z - 1 & \text{if } z \geq \frac{g(s_n|\theta_k)}{g(s_n|\theta_i)}. \end{cases}$$

The inverse hazard rate $h^f_\sigma(\theta_i|s_j)$ is increasing in j under MLRP as

$$h^f_\sigma(\theta_i|s_j) = \sum_{k=i+1}^m \frac{g^f_\sigma(\theta_k|s_j)}{g^f_\sigma(\theta_i|s_j)} = \sum_{k=i+1}^m \frac{g(s_j|\theta_k) f(\theta_k)}{g(s_j|\theta_i) f(\theta_i)}$$

and each element of the sum is increasing under MLRP. Thus, under MLRP

$$M^f_\sigma(y|\theta_i) = \begin{cases} 0 & \text{if } z < \sum_{k=i+1}^m \frac{f(\theta_k)}{f(\theta_i)} \frac{g(s_1|\theta_k)}{g(s_1|\theta_i)}, \\ G(s_j|\theta_i) & \text{if } z \in \left[\sum_{k=i+1}^m \frac{f(\theta_k)}{f(\theta_i)} \frac{g(s_j|\theta_k)}{g(s_j|\theta_i)}, \sum_{k=i+1}^m \frac{f(\theta_k)}{f(\theta_i)} \frac{g(s_{j+1}|\theta_k)}{g(s_{j+1}|\theta_i)}\right) \text{ and } 1 \leq j < n, \\ 1 & \text{if } z \geq \sum_{k=i+1}^m \frac{f(\theta_k)}{f(\theta_i)} \frac{g(s_n|\theta_k)}{g(s_n|\theta_i)}. \end{cases}$$

Hence,

$$\int_0^z M^f_\sigma(y|\theta_i) dy = \begin{cases} 0 & \text{if } z < \sum_{k=i+1}^m \frac{f(\theta_k)}{f(\theta_i)} \frac{g(s_1|\theta_k)}{g(s_1|\theta_i)}, \\ G(s_j|\theta_i) z - \sum_{k=i+1}^m \frac{f(\theta_k)}{f(\theta_i)} G(s_j|\theta_k) & \text{if } z \in \left[\sum_{k=i+1}^m \frac{f(\theta_k)}{f(\theta_i)} \frac{g(s_j|\theta_k)}{g(s_j|\theta_i)}, \sum_{k=i+1}^m \frac{f(\theta_k)}{f(\theta_i)} \frac{g(s_{j+1}|\theta_k)}{g(s_{j+1}|\theta_i)}\right) \\ & \quad \text{and } 1 \leq j < n, \\ z - \sum_{k=i+1}^m \frac{f(\theta_k)}{f(\theta_i)} & \text{if } z \geq \sum_{k=i+1}^m \frac{f(\theta_k)}{f(\theta_i)} \frac{g(s_n|\theta_k)}{g(s_n|\theta_i)}. \end{cases}$$

To show that $\breve{\sigma} \succeq^f_{HRS} \hat{\sigma} \;\forall f \implies \breve{\sigma} \succeq_{LK} \hat{\sigma}$, I prove the contrapositive statement $\breve{\sigma} \not\succeq_{LK} \hat{\sigma} \implies \exists f,\; \breve{\sigma} \not\succeq^f_{HRS} \hat{\sigma}$. If $\breve{\sigma} \not\succeq_{LK} \hat{\sigma}$, there exist $\theta_i, \theta_k \in \Theta$ with $i < k$ and $\tilde{z} > 0$ such that

$$\int_0^{\tilde{z}} L_{\breve{\sigma}}(y|\theta_i, \theta_k) dy < \int_0^{\tilde{z}} L_{\hat{\sigma}}(y|\theta_i, \theta_k) dy.$$

	θ_1	θ_2	θ_3
$\check{s}_1$	0.5	0.4	0.17
$\check{s}_2$	0.33	0.33	0.33
$\check{s}_3$	0.17	0.27	0.5

$\check{g}(\check{s}|\theta)$

	θ_1	θ_2	θ_3
$\hat{s}_1$	0.42	0.41	0.17
$\hat{s}_2$	0.33	0.33	0.33
$\hat{s}_3$	0.25	0.26	0.5

$\hat{g}(\hat{s}|\theta)$

Fig. 1. Two signals $\check{\sigma} = (\{\check{s}_1, \check{s}_2, \check{s}_3\}, \check{g})$ and $\hat{\sigma} = (\{\hat{s}_1, \hat{s}_2, \hat{s}_3\}, \hat{g})$.

	θ_1	θ_2	θ_3
$z \in [0, 0.387)$	0.1	0.9	1
$z \in [0.387, 0.717)$	0.082	0.865	1
$z \in [0.717, 1)$	0.060	0.823	0.5

$\Pr(\theta \le x|\check{G}^f(s) > z)$

	θ_1	θ_2	θ_3
$z \in [0, 0.387)$	0.1	0.9	1
$z \in [0.387, 0.717)$	0.095	0.865	1
$z \in [0.717, 1)$	0.088	0.823	1

$\Pr(\theta \le x|\hat{G}^f(s) > z)$

Fig. 2. CDFs of random variables $\theta|\check{G}^f(s) > z$ and $\theta|\hat{G}^f(s) > z$ for $z \in [0, 1)$.

Define the prior f by $f(\theta_\kappa) = \prod_{l=1}^{\kappa} \zeta_l$ for all $\kappa \neq k$ where $\zeta_l \in [0, 1]$ for all $l \in \{1, \dots, m\}$, and $f(\theta_k) = 1 - \sum_{\kappa \neq j} \prod_{l=1}^{\kappa} \zeta_l$. For this prior, we have for any signal σ and any $z \ge 0$

$$\lim_{\zeta_{i+1} \to 0} \int_0^z M^f_\sigma(y|\theta_i)dy = \int_0^z L_\sigma\left(\frac{f(\theta_i)}{f(\theta_k)} y|\theta_i, \theta_k\right) dy = \frac{f(\theta_k)}{f(\theta_i)} \int_0^{\frac{f(\theta_k)}{f(\theta_i)} z} L_\sigma(x|\theta_i, \theta_k)dx.$$

Thus, $\check{\sigma} \not\succeq^f_{HRS} \hat{\sigma}$ for the prior specified above with ζ_{i+1} being small as

$$\lim_{\zeta_{i+1} \to 0} \int_0^{\frac{f(\theta_i)}{f(\theta_k)} \tilde{z}} M^f_{\check{\sigma}}(y|\theta_i)dy < \lim_{\zeta_{i+1} \to 0} \int_0^{\frac{f(\theta_i)}{f(\theta_k)} \tilde{z}} M^f_{\hat{\sigma}}(y|\theta_i)dy.$$

Thus, 2) implies 1) for any signals $\check{\sigma}$ and $\hat{\sigma}$ that satisfy MLRP.

It remains to prove that 3) implies 2). We prove the contrapositive statement. Suppose that 2) is not satisfied for two signals $\check{\sigma}$ and $\hat{\sigma}$. Then there exists a prior $\tilde{f}$, a type $\tilde{\theta} \in \Theta$, and a number $\tilde{z} > 0$ such that

$$\int_0^{\tilde{z}} M^{\tilde{f}}_{\check{\sigma}}(z|\tilde{\theta})dz < \int_0^{\tilde{z}} M^{\tilde{f}}_{\hat{\sigma}}(z|\tilde{\theta})dz.$$

We now use the proof of necessity of the HRS order in Theorem 2 in the same way as in the proof of Proposition 1 to show that 3) cannot hold. It follows that 3) implies 2). □

A.7. Proof of Proposition 4

I prove the proposition with two examples. Suppose $\Theta = \{\theta_1, \theta_2, \theta_3\}$. Set at first the prior to $f(\theta_1) = 0.1$, $f(\theta_2) = 0.8$, and $f(\theta_3) = 0.1$. Consider now the signals defined in Fig. 1.

I aim to prove that $\check{\sigma} \succeq^f_{MIO-ND} \hat{\sigma}$ but $\check{\sigma} \not\succeq^f_{HRS} \hat{\sigma}$. Let $\check{G}^f(\check{s})$ and $\hat{G}^f(\hat{s})$ denote the marginal distributions of the signals $\check{\sigma}$ and $\hat{\sigma}$ for the prior f. By Proposition 3 in Athey and Levin (2018), we have $\check{\sigma} \succeq^f_{MIO-ND} \hat{\sigma}$ if and only if the random variable $\theta|\check{G}^f(s) > z$ first-order stochastically dominates the random variable $\theta|\hat{G}^f(s) > z$ for all $z \in [0, 1)$. Fig. 2 depicts the conditional cumulative distribution functions for these random variables.

One can deduce from these tables that $\Pr(\theta \le x|\hat{G}^f(s) > z) \le \Pr(\theta \le x|\check{G}^f(s) > z)$ for all $x \in \Theta$ and $z \in [0, 1)$. Thus, $\check{\sigma} \succeq^f_{MIO-ND} \hat{\sigma}$.

Next, one can compute the following inverse hazard rates of $\check{\sigma}$ and $\hat{\sigma}$ (Fig. 3).

As $\check{h}^f(\theta_2|\check{s}_1) > \hat{h}^f(\theta_2|\hat{s}_1)$, it follows that $\check{M}^f(z|\theta_2)$ cannot be a mean-preserving spread of $\hat{M}^f(z|\theta_2)$. Thus, $\check{\sigma} \not\succeq^f_{HRS} \hat{\sigma}$.

This establishes that MIO-ND does not imply the HRS order. As MIO-ND implies IPO by Theorem 1(i) in Ganuza and Penalva (2010), it follows that also IPO does not imply the HRS order.

The following example proves that the reverse is also true: the HRS order does neither imply IPO nor MIO-ND. Let $\theta_1 = 1$, $\theta_2 = 2$, and $\theta_3 = 4$ and set the prior to $f'(\theta_1) = \frac{112}{140}$, $f'(\theta_2) = \frac{27}{140}$, and $f'(\theta_3) = \frac{1}{140}$. Consider the signals $\check{\sigma}'$ and $\hat{\sigma}'$ given in Fig. 4.

	$\check{h}^f(\theta_1\|\check{s})$	$\check{h}^f(\theta_2\|\check{s})$
$\check{s}_1$	6.74	0.053
$\check{s}_2$	9	0.125
$\check{s}_3$	15.647	0.231

	$\hat{h}^f(\theta_1\|\hat{s})$	$\hat{h}^f(\theta_2\|\hat{s})$
$\hat{s}_1$	8.214	0.052
$\hat{s}_2$	9	0.125
$\hat{s}_3$	10.32	0.240

Fig. 3. Inverse hazard rates for $\check{\sigma}$ and $\hat{\sigma}$.

	θ_1	θ_2	θ_3
$\check{s}_1$	$\frac{1}{2}$	$\frac{8}{15}$	$\frac{1}{5}$
$\check{s}_2$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$\check{s}_3$	$\frac{1}{6}$	$\frac{2}{15}$	$\frac{7}{15}$

$\check{g}'(\check{s}|\theta)$

	θ_1	θ_2	θ_3
$\hat{s}_1$	$\frac{5}{12}$	$\frac{5}{12}$	$\frac{1}{4}$
$\hat{s}_2$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$\hat{s}_3$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{5}{12}$

$\hat{g}'(\hat{s}|\theta)$

Fig. 4. Two signals $\check{\sigma}' = (\{\check{s}_1, \check{s}_2, \check{s}_3\}, \check{g}')$ and $\hat{\sigma}' = (\{\hat{s}_1, \hat{s}_2, \hat{s}_3\}, \hat{g}')$.

	$\check{h}'^{f'}(\theta_1\|\check{s})$	$\check{h}'^{f'}(\theta_2\|\check{s})$
$\check{s}_1$	0.261	0.014
$\check{s}_2$	0.25	0.037
$\check{s}_3$	0.218	0.127

	$\hat{h}'^{f'}(\theta_1\|\hat{s})$	$\hat{h}'^{f'}(\theta_2\|\hat{s})$
$\hat{s}_1$	0.246	0.022
$\hat{s}_2$	0.25	0.037
$\hat{s}_3$	0.256	0.061

Fig. 5. Inverse hazard rates for $\check{\sigma}'$ and $\hat{\sigma}'$.

	$E[\theta\|\check{s}_i]$	$E[\theta\|\hat{s}_i]$
$i=1$	1.212	1.206
$i=2$	1.214	1.214
$i=3$	1.220	1.227

Fig. 6. Inverse hazard rates for $\check{\sigma}'$ and $\hat{\sigma}'$.

The inverse hazard rates for these signals can be calculated as shown in Fig. 5. one can quickly verify that for all given θ, the distribution over the values in $\{\hat{h}'^{f'}(\theta|\hat{s}_i)\}_{i=1,2,3}$ can be obtained from mixing the values in $\{\check{h}'^{f'}(\theta|\check{s}_i)\}_{i=1,2,3}$. Thus, $\check{\sigma}' \succeq^{f'}_{HRS} \hat{\sigma}'$.

For $\check{\sigma}' \succeq^{f'}_{IPO} \hat{\sigma}'$ to hold, the distribution over posterior means induced by $\check{\sigma}'$ would need to be a mean-preserving spread of the distribution of posterior means induced by $\hat{\sigma}'$. The posterior means can be computed as shown in Fig. 6. Clearly, we have $\check{\sigma}' \not\succeq^{f'}_{IPO} \hat{\sigma}'$ as $E[\theta|\hat{s}_1] < E[\theta|\check{s}_1]$.

Thus, the HRS order does not imply IPO, and therefore also not MIO-ND, as MIO-ND implies IPO. $\square$

A.8. Proof of Lemma 2

The binding constraints imply that under any solution $\{(q^*(\cdot|s), t^*(\cdot|s))\}_{s\in S}$ to ($\mathbf{P}^f_\sigma(X, u^P, u^A)$), we have $v^A(\theta_1, q^*(\theta_1|s)) - t^*(\theta_1|s) = 0$ and

$$v^A(\theta_i, q^*(\theta_i|s)) - t^*(\theta_i|s) = \sum_{k=1}^{i-1} \Delta^A_v(\theta_k, q^*(\theta_k|s)) \quad \forall i = \{2, \dots, m\}.$$

Basic manipulations yield

$$\begin{aligned}\sum_{i=1}^{m}(v^A(\theta_i, q^*(\theta_i|s)) - t^*(\theta_i|s))g^f_\sigma(\theta_i|s_j) &= \sum_{i=1}^{m} \Delta^A_v(\theta_i, q^*(\theta_i|s_j))(1 - G^f_\sigma(\theta_i|s_j)) \\ &= \sum_{i=1}^{m} h^f_\sigma(\theta_i|s_j)\Delta^A_v(\theta_i, q^*(\theta_i|s_j))g^f_\sigma(\theta_i|s_j).\end{aligned}$$

Thus, the ex-ante welfare can be rewritten as

$$W^f_\sigma(X, v^P, v^A, v^T)$$

$$
\begin{aligned}
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\Big(\beta^P \bar{v}(\theta_i, q^*(\theta_i|s_j)) - (\beta^P - \beta^A) h_\sigma^f(\theta_i|s_j)\Delta_v^A(\theta_i, q^*(\theta_i|s_j)) \\
&\qquad + \beta^T v^T(\theta_i, q^*(\theta_i|s_j))\Big) g(s_j|\theta_i) f(\theta_i) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} \omega(\theta_i, h_\sigma^f(\theta_i|s_j)) g(s_j|\theta_i) f(\theta_i) = \sum_{i=1}^{m} f(\theta_i)\int_0^\infty \omega(\theta_i, z)\, dM_\sigma^f(z|\theta_i),
\end{aligned}
$$

where the second equality follows under regularity from the definition of the function $\omega(\theta_i, z)$ and the third equality is implied by the same change of variable as in the proof of Lemma 1. □

A.9. Proof of Theorem 3

For $\check{\sigma} \succeq_{HRS}^f \hat{\sigma}$, $M_{\check{\sigma}}^f(z|\theta)$ is a mean-preserving spread of $M_{\hat{\sigma}}^f(z|\theta)$ and it follows directly from Theorem 3.A.1 in Shaked and Shanthikumar (2007) that for each $\theta \in \Theta$

$$
\int \omega(\theta, z)\, dM_{\check{\sigma}}^f(z|\theta) \geq \int \omega(\theta, z)\, dM_{\hat{\sigma}}^f(z|\theta)
$$

if $\omega(\theta, \cdot)$ is convex for all $\theta \in \Theta$ whereas the converse is true if $\omega(\theta, \cdot)$ is concave for all $\theta \in \Theta$. □

A.10. Proof of Proposition 5

It is shown in the proof of Proposition 1 that under the conditions stated in the Proposition, a decision is optimal in the relaxed problem ($\mathbf{R}_\sigma^f(X, u^P, u^A)$) if and only if it is supported on

$$
X_\sigma^{*,f}(\theta_i|s_j) \equiv \arg\max_{x\in X} \bar{u}(\theta_i, x) - h_\sigma^f(\theta_i|s_j)\Delta_u^A(\theta_i, x).
$$

Moreover, it is argued that due to the strict supermodularity of the deterministic virtual surplus function, any selection from $X_\sigma^{*,f}(\theta_i|s_j)$ is non-decreasing in θ_i by Topkis' monotonicity theorem. Thus, any welfare-optimal selection is also non-decreasing and can therefore be implemented under the original problem ($\mathbf{P}_\sigma^f(X, u^P, u^A)$), too. □

A.11. Proof of Proposition 6

Proposition 5 implies that $(X, u^P, u^A, u^T) = ([0, \bar{x}], -c(x), \phi(x) + \theta x, 0) \in \mathcal{U}_{\check{\sigma}}^{\prime R,f} \cap \mathcal{U}_{\hat{\sigma}}^{\prime R,f}$ as $u^A(\theta, x) = \phi(x) + \theta x$ is strictly supermodular and increasing in θ and $-\Delta_u^A(\theta_i, x) = -\Delta\theta x$ is supermodular. I can therefore apply Theorem 3. Total surplus is reflected by the welfare weights $\beta^P = \beta^A = 1$. As $\phi(x) - c(x)$ is strictly concave, the function

$$
\rho(\theta_i, z) = \arg\max_x \phi(x) + (\theta_i - z\Delta\theta)x - c(x)
$$

is well-defined. It therefore remains to show that the function

$$
\omega(\theta_i, z) = \phi(\rho(\theta_i, z)) + \theta\rho(\theta_i, z) - c(\rho(\theta_i, z)
$$

is concave. Using the condition of strictly positive production, I obtain

$$
\frac{\partial^2 \omega(\theta_i, z)}{\partial z^2} = \Delta\theta^2 \left.\frac{(\phi''(x) - c''(x))^2 - (\phi'(x) + \theta_i - c'(x))(\phi'''(x) - c'''(x))}{(\phi''(x) - c''(x))^3}\right|_{x=\rho(\theta_i, z)}.
$$

By the strict concavity of $\phi(x) - c(x)$, this term is negative if $\phi'(x) + \theta_i - c'(x)$ is log-concave for any $\theta_i \in \Theta$. □

A.12. Proof of Proposition 7

Proposition 5 implies that $(X, u^P, u^A, u^T) = ([0, \bar{x}], 0, R(x) - (\kappa - \theta)x, \int_0^x (P(y) - P(x))dx) \in \mathcal{U}_{\check{\sigma}}^{\prime R,f} \cap \mathcal{U}_{\hat{\sigma}}^{\prime R,f}$ as $u^A(\theta, x) = R(x) - (\kappa - \theta)x$ is strictly supermodular and increasing in θ and $-\Delta_u^A(\theta_i, x) = -\Delta\theta x$ is supermodular. I apply Theorem 3 for the welfare weights $\beta^P = \beta^A = \beta^T = 1$ that reflect total surplus. The function

$$
\rho(\theta_i, z) = \arg\max_x R(x) - (\kappa - \theta_i + z\Delta\theta)x
$$

is well-defined due to strict concavity of $R(x)$. It remains to prove that the function

$$
\omega(\theta_i, z) = \int_0^{\rho(\theta_i, z)} (P(x) - (\kappa - \theta_i))dx
$$

	θ_1	θ_2	θ_3
$\check{s}_1$	$^1/_2$	$^{(1+\epsilon)}/_3$	$^1/_5$
$\check{s}_2$	$^1/_3$	$^1/_3$	$^1/_3$
$\check{s}_3$	$^1/_6$	$^{(1-\epsilon)}/_3$	$^7/_{15}$

$\check{g}(\check{s}|\theta)$

	θ_1	θ_2	θ_3
$\hat{s}_1$	$^5/_{12}$	$^1/_3$	$^1/_4$
$\hat{s}_2$	$^1/_3$	$^1/_3$	$^1/_3$
$\hat{s}_3$	$^1/_4$	$^1/_3$	$^5/_{12}$

$\hat{g}(\hat{s}|\theta)$

Fig. 7. Two signals $\check{\sigma} = (\{\check{s}_1, \check{s}_2, \check{s}_3\}, \check{g})$ and $\hat{\sigma} = (\{\hat{s}_1, \hat{s}_2, \hat{s}_3\}, \hat{g})$.

is concave in z for all θ_i. Under strictly positive production, I obtain

$$\frac{\partial^2 \omega(\theta_i, z)}{\partial z^2} = \Delta\theta^2 \left.\frac{P'(x)R''(x) - (P(x) - (\kappa - \theta_i))R'''(x)}{R''(x)^3}\right|_{x=\rho(\theta_i,z)}$$

By the strict concavity of $R(x)$, this expression is negative if marginal revenue is concave, i.e., $R'''(x) \le 0$. □

A.13. Proof of Proposition 8

I have shown in the proof of Proposition 7 that $(X, u^P, u^A, u^T) = ([0, \bar{x}], 0, R(x) - (\kappa - \theta)x, \int_0^x (P(y) - P(x))dx) \in \mathcal{U}'^{R,f}_{\check{\sigma}} \cap \mathcal{U}'^{R,f}_{\check{\sigma}}$ and that the function

$$\rho(\theta_i, z) = \arg\max_x R(x) - (\kappa - \theta_i + z\Delta\theta)x$$

is well-defined. To analyze the consumer surplus effects, I apply Theorem 3 for the welfare weights $\beta^P = \beta^A = 0$ and $\beta^T = 1$. Using the definition $CS(x) \equiv \int_0^x (P(y) - P(x))dy$, I want to show that the function

$$\omega(\theta_i, z) = CS(\rho(\theta_i, z))$$

is convex in z for all θ_i. It is straightforward to check that $\rho(\theta_i, z)$ and $\omega(\theta_i, z)$ are weakly decreasing in z and are zero for large z. For $\rho(\theta_i, z) > 0$, I obtain

$$\frac{\partial^2 \omega(\theta_i, z)}{\partial z^2} = \Delta\theta^2 \left.\frac{CS''(x)R''(x) - CS'(x)R'''(x)}{R''(x)^3}\right|_{x=\rho(\theta_i,z)}$$

Using the definitions of the functions $CS(x)$ and $R(x)$, the term above is positive if

$$\frac{P'(x) + xP''(x)}{xP'(x)} \ge \frac{3P''(x) + xP'''(x)}{2P'(x) + xP''(x)}.$$

Due to $P'(x) < 0$ and $R''(x) < 0$, this is equivalent to the condition

$$\log(-P'(x))'' = \frac{P'(x)P'''(x) - P''(x)^2}{P'(x)^2} \le \frac{2}{x^2}.$$

Thus, the function $\omega(\theta_i, z)$ is convex and positive for small values z if $R(x)$ is strictly concave and $[\log(-P'(x))'' \le \frac{2}{x^2}$, and takes the value zero for large values of z. The function is therefore globally convex in z for any θ_i. □

Appendix B. Relation between information orders: example

In this part of the appendix, I illustrate the relation between the HRS order and other information order by the use of examples. Consider first two signals $\check{\sigma}$ and $\hat{\sigma}$ as defined in the following figure for the parameter $\epsilon \in [0, 1)$.

Lemma 6. *The signals $\check{\sigma}$ and $\hat{\sigma}$ defined in Fig. 7 are not Blackwell ordered.*

Proof. It needs to be shown that there does not exist a matrix B that satisfies the conditions of Definition 1 with respect to $\check{\sigma}$ and $\hat{\sigma}$. Note first that $\hat{\sigma} \succeq_B \check{\sigma}$ is not possible as the range of $\{\check{g}(s|\theta_i)\}_S$ contains the range of $\{\hat{g}(s|\theta_i)\}_S$ for any $\theta_i \in \{\theta_1, \theta_2, \theta_3\}$. For $\check{\sigma} \succeq_B \hat{\sigma}$ to hold, a matrix B would need to satisfy

$$\frac{1}{2}b_{11} + \frac{1}{3}b_{21} + \frac{1}{6}b_{31} = \frac{5}{12}, \tag{1}$$

$$\frac{1+\epsilon}{3}b_{11} + \frac{1}{3}b_{21} + \frac{1-\epsilon}{3}b_{31} = \frac{1}{3}, \tag{2}$$

$$\frac{1}{5}b_{11} + \frac{1}{3}b_{21} + \frac{7}{15}b_{31} = \frac{1}{4}. \tag{3}$$

	$\frac{\check{g}(\check{s}\mid\theta_2)}{\check{g}(\check{s}\mid\theta_1)}$	$\frac{\check{g}(\check{s}\mid\theta_3)}{\check{g}(\check{s}\mid\theta_1)}$	$\frac{\check{g}(\check{s}\mid\theta_3)}{\check{g}(\check{s}\mid\theta_2)}$
$\check{s}_1$	$\frac{2(1+\varepsilon)}{3}$	$\frac{2}{5}$	$\frac{3}{5(1+\varepsilon)}$
$\check{s}_2$	1	1	1
$\check{s}_3$	$2(1-\varepsilon)$	$\frac{14}{5}$	$\frac{7}{5(1-\varepsilon)}$

	$\frac{\hat{g}(s\mid\theta_2)}{\hat{g}(s\mid\theta_1)}$	$\frac{\hat{g}(s\mid\theta_3)}{\hat{g}(s\mid\theta_1)}$	$\frac{\hat{g}(s\mid\theta_3)}{\hat{g}(s\mid\theta_2)}$
$\hat{s}_1$	$^4/_5$	$^3/_5$	$^3/_4$
$\hat{s}_2$	1	1	1
$\hat{s}_3$	$^4/_3$	$^5/_3$	$^5/_4$

Fig. 8. Likelihood ratios of $\check{\sigma}$ and $\hat{\sigma}$.

Taking the differences (1)-(2) and (1)-(3) yields

$$b_{11} - b_{31} = \frac{1}{2-4\varepsilon},$$
$$b_{11} - b_{31} = \frac{5}{9},$$

which is only possible for $\varepsilon = \frac{1}{20}$. Setting $\varepsilon = \frac{1}{20}$, I use $b_{11} = b_{31} + \frac{5}{9}$ in (2) to obtain $b_{21} = -2b_{31} + \frac{5}{12}$. This yields a contradiction as

$$1 = b_{11} + b_{21} + b_{31} = (b_{31} + \tfrac{5}{9}) + (-2b_{31} + \tfrac{5}{12}) + b_{31} = \frac{35}{36} \neq 1.$$

Thus, $\check{\sigma} \not\succeq_B \hat{\sigma}$. $\square$

Lemma 7. *The signals $\check{\sigma}$ and $\hat{\sigma}$ defined in Fig. 7 have monotone likelihood ratios if and only if $\varepsilon \leq \frac{1}{2}$.*

Proof. The likelihood ratios of the two signals can be computed as shown in Fig. 8.

The likelihood ratios of $\hat{\sigma}$ are monotone. The likelihood ratios of $\check{\sigma}$ are monotone for $\frac{2(1+\varepsilon)}{3} \leq 1 \leq 2(1-\varepsilon) \iff \varepsilon \leq \frac{1}{2}$ and $\frac{3}{5(1+\varepsilon)} \leq 1 \leq \frac{7}{5(1-\varepsilon)} \impliedby \varepsilon \geq 0$. $\square$

Lemma 8. *The signals $\check{\sigma}$ and $\hat{\sigma}$ defined in Fig. 7 satisfy $\check{\sigma} \succeq_{LK} \hat{\sigma}$ for $\varepsilon \leq \frac{1}{5}$.*

Proof. Using the likelihood ratios of the two signals in Fig. 8, one can compute

$$\check{L}(z|\theta_2,\theta_1) = \begin{cases} 0 & \text{if} \quad z < \frac{2(1+\varepsilon)}{3}, \\ \frac{1}{2} & \text{if} \quad z \in \left[\frac{2(1+\varepsilon)}{3}, 1\right), \\ \frac{5}{6} & \text{if} \quad z \in [1, 2(1-\varepsilon)), \\ 1 & \text{if} \quad z \geq 2(1-\varepsilon), \end{cases} \qquad \check{L}(z|\theta_3,\theta_1) = \begin{cases} 0 & \text{if} \quad z < \frac{2}{5}, \\ \frac{1}{2} & \text{if} \quad z \in \left[\frac{2}{5}, 1\right), \\ \frac{5}{6} & \text{if} \quad z \in \left[1, \frac{14}{5}\right), \\ 1 & \text{if} \quad z \geq \frac{14}{5}, \end{cases}$$

$$\check{L}(z|\theta_3,\theta_2) = \begin{cases} 0 & \text{if} \quad z < \frac{3}{5(1+\varepsilon)}, \\ \frac{1+\varepsilon}{3} & \text{if} \quad z \in \left[\frac{3}{5(1+\varepsilon)}, 1\right), \\ \frac{2+\varepsilon}{3} & \text{if} \quad z \in \left[1, \frac{7}{5(1-\varepsilon)}\right), \\ 1 & \text{if} \quad z \geq \frac{7}{5(1-\varepsilon)}, \end{cases}$$

and

$$\hat{L}(z|\theta_2,\theta_1) = \begin{cases} 0 & \text{if} \quad z < \frac{4}{5}, \\ \frac{5}{12} & \text{if} \quad z \in \left[\frac{4}{5}, 1\right), \\ \frac{3}{4} & \text{if} \quad z \in \left[1, \frac{4}{3}\right), \\ 1 & \text{if} \quad z \geq \frac{4}{3}, \end{cases} \qquad \hat{L}(z|\theta_3,\theta_1) = \begin{cases} 0 & \text{if} \quad z < \frac{3}{5}, \\ \frac{5}{12} & \text{if} \quad z \in \left[\frac{3}{5}, 1\right), \\ \frac{3}{4} & \text{if} \quad z \in \left[1, \frac{5}{3}\right), \\ 1 & \text{if} \quad z \geq \frac{5}{3}. \end{cases}$$

$$\hat{L}(z|\theta_3,\theta_2) = \begin{cases} 0 & \text{if} \quad z < \frac{3}{4}, \\ \frac{1}{3} & \text{if} \quad z \in \left[\frac{3}{4}, 1\right), \\ \frac{2}{3} & \text{if} \quad z \in \left[1, \frac{5}{4}\right), \\ 1 & \text{if} \quad z \geq \frac{5}{4}. \end{cases}$$

	$\check{h}(\theta_1\vert\check{s})$	$\check{h}(\theta_2\vert\check{s})$
$\check{s}_1$	$\frac{6f(\theta_3)+10(1+\varepsilon)f(\theta_2)}{15f(\theta_1)}$	$\frac{3f(\theta_3)}{5(1+\varepsilon)f(\theta_2)}$
$\check{s}_2$	$\frac{f(\theta_3)+f(\theta_2)}{f(\theta_1)}$	$\frac{f(\theta_3)}{f(\theta_2)}$
$\check{s}_3$	$\frac{42f(\theta_3)+30(1-\varepsilon)f(\theta_2)}{15f(\theta_1)}$	$\frac{7f(\theta_3)}{5(1-\varepsilon)f(\theta_2)}$

	$\hat{h}(\theta_1\vert\hat{s})$	$\hat{h}(\theta_2\vert\hat{s})$
$\hat{s}_1$	$\frac{9f(\theta_3)+12f(\theta_2)}{15f(\theta_1)}$	$\frac{3f(\theta_3)}{4f(\theta_2)}$
$\hat{s}_2$	$\frac{f(\theta_3)+f(\theta_2)}{f(\theta_1)}$	$\frac{f(\theta_3)}{f(\theta_2)}$
$\hat{s}_3$	$\frac{25f(\theta_3)+20f(\theta_2)}{15f(\theta_1)}$	$\frac{5f(\theta_3)}{4f(\theta_2)}$

Fig. 9. Inverse hazard rates for $\check{\sigma}$ and $\hat{\sigma}$.

Given two CDFs $\check{L}(z)$ and $\hat{L}(z)$ with identical mean, $\check{L}(z)$ is a mean-preserving spread of $\hat{L}(z)$ if there exists some $Z>0$ such that $\check{L}(z)\geq \hat{L}(z)$ for all $z<Z$ and $\check{L}(z)\leq\hat{L}(z)$ for all $z>Z$. For $\check{L}(z|\theta_3,\theta_1)$ and $\hat{L}(z|\theta_3,\theta_1)$, we have $Z=\frac{5}{3}$. For $\check{L}(z|\theta_3,\theta_2)$ and $\hat{L}(z|\theta_3,\theta_2)$, we have $Z=\frac{7}{5(1-\varepsilon)}$ as $\frac{7}{5(1-\varepsilon)}>\frac{5}{4}$. Finally, for $\check{L}(z|\theta_2,\theta_1)$ and $\hat{L}(z|\theta_2,\theta_1)$, we have $Z=2(1-\varepsilon)$ for $\varepsilon\leq\frac{1}{5}\iff\frac{3}{5(1+\varepsilon)}\leq\frac{3}{4}\wedge 2(1-\varepsilon)\geq\frac{5}{4}$. □

Lemma 9. *The signals $\check{\sigma}$ and $\hat{\sigma}$ defined in Fig. 7 satisfy $\check{\sigma}\succeq^f_{HRS}\hat{\sigma}$ for any prior f on $\{\theta_1,\theta_2,\theta_3\}$ if $\varepsilon\leq\frac{1}{5}$. For $\varepsilon>\frac{1}{5}$, $\check{\sigma}\succeq^f_{HRS}\hat{\sigma}$ if the prior f satisfies $\frac{f(\theta_3)}{f(\theta_2)}\geq\frac{10\varepsilon-2}{3}$.*

Proof. The inverse hazard rates of the two signals can be computed as shown in Fig. 9.

Using these inverse hazard rates, one can compute

$$\check{M}^f_\sigma(z|\theta_2)=\begin{cases}0 & \text{if}\quad z\in\left[0,\frac{3f(\theta_3)}{5(1+\varepsilon)f(\theta_2)}\right),\\ \frac{1+\varepsilon}{3} & \text{if}\quad z\in\left[\frac{3f(\theta_3)}{5(1+\varepsilon)f(\theta_2)},\frac{f(\theta_3)}{f(\theta_2)}\right),\\ \frac{2+\varepsilon}{3} & \text{if}\quad z\in\left[\frac{f(\theta_3)}{f(\theta_2)},\frac{7f(\theta_3)}{5(1-\varepsilon)f(\theta_2)}\right),\\ 1 & \text{if}\quad z\geq\frac{7f(\theta_3)}{5(1-\varepsilon)f(\theta_2)},\end{cases}\quad\text{and}\quad \hat{M}^f_\sigma(z|\theta_2)=\begin{cases}0 & \text{if}\quad z\in\left[0,\frac{3f(\theta_3)}{4f(\theta_2)}\right),\\ \frac{1}{3} & \text{if}\quad z\in\left[\frac{3f(\theta_3)}{4f(\theta_2)},\frac{f(\theta_3)}{f(\theta_2)}\right),\\ \frac{2}{3} & \text{if}\quad z\in\left[\frac{f(\theta_3)}{f(\theta_2)},\frac{5f(\theta_3)}{4f(\theta_2)}\right),\\ 1 & \text{if}\quad z\geq\frac{5f(\theta_3)}{4f(\theta_2)}.\end{cases}$$

Note that $\check{L}(z|\theta_3,\theta_2)$ being a mean-preserving spread of $\hat{L}(z|\theta_3,\theta_2)$ implies that $\check{M}^f_\sigma(z|\theta_2)$ is a mean-preserving spread of $\hat{M}^f_\sigma(z|\theta_2)$.

Further, one can compute

$$\check{M}^f_\sigma(z|\theta_1)=\begin{cases}0 & \text{if}\quad z\in\left[0,\frac{6f(\theta_3)+10(1+\varepsilon)f(\theta_2)}{15f(\theta_1)}\right),\\ \frac{1}{2} & \text{if}\quad z\in\left[\frac{6f(\theta_3)+10(1+\varepsilon)f(\theta_2)}{15f(\theta_1)},\frac{f(\theta_3)+f(\theta_2)}{f(\theta_1)}\right),\\ \frac{5}{6} & \text{if}\quad z\in\left[\frac{f(\theta_3)+f(\theta_2)}{f(\theta_1)},\frac{42f(\theta_3)+30(1-\varepsilon)f(\theta_2)}{15f(\theta_1)}\right),\\ 1 & \text{if}\quad z\geq\frac{42f(\theta_3)+30(1-\varepsilon)f(\theta_2)}{15f(\theta_1)},\end{cases}$$

and

$$\hat{M}^f_\sigma(z|\theta_1)=\begin{cases}0 & \text{if}\quad z\in\left[0,\frac{9f(\theta_3)+12f(\theta_2)}{15f(\theta_1)}\right),\\ \frac{5}{12} & \text{if}\quad z\in\left[\frac{9f(\theta_3)+12f(\theta_2)}{15f(\theta_1)},\frac{f(\theta_3)+f(\theta_2)}{f(\theta_1)}\right),\\ \frac{3}{4} & \text{if}\quad z\in\left[\frac{f(\theta_3)+f(\theta_2)}{f(\theta_1)},\frac{25f(\theta_3)+20f(\theta_2)}{15f(\theta_1)}\right),\\ 1 & \text{if}\quad z\geq\frac{25f(\theta_3)+20f(\theta_2)}{15f(\theta_1)}.\end{cases}$$

For $\frac{6f(\theta_3)+10(1+\varepsilon)f(\theta_2)}{15f(\theta_1)}\leq\frac{9f(\theta_3)+12f(\theta_2)}{15f(\theta_1)}<\frac{f(\theta_3)+f(\theta_2)}{f(\theta_1)}\iff\frac{f(\theta_3)}{f(\theta_2)}\geq\frac{10\varepsilon-2}{3}$, we have that $\check{M}^f_\sigma(z|\theta_1)\geq\hat{M}^f_\sigma(z|\theta_1)$ for $z<\frac{25f(\theta_3)+20f(\theta_2)}{15f(\theta_1)}$ and $\check{M}^f_\sigma(z|\theta_1)\leq\hat{M}^f_\sigma(z|\theta_1)$ for $z>\frac{25f(\theta_3)+20f(\theta_2)}{15f(\theta_1)}$. Thus, $M^f_\sigma(z|\theta_1)$ is a mean-preserving spread of $\hat{M}^f_\sigma(z|\theta_1)$ for $\frac{f(\theta_3)}{f(\theta_2)}\geq\frac{10\varepsilon-2}{3}$. Note that this last condition is satisfied for any prior if and only if $\varepsilon\leq\frac{1}{5}$. □

Appendix C. Relation to Haghpanah and Siegel (2023)

In this appendix, I use a variant from the example of Section 2.5 to illustrate the relationship of Proposition 6 to the result in Haghpanah and Siegel (2023). In particular, I want to show that the difference between a continuous and a discrete set of production possibilities X can drive the difference in results. As before, the agent receives gross utility of θx from consuming $x\in[0,1]$ units of the good where θ is uniformly distributed on $\{\theta_1,\theta_2\}$ with $0<\theta_1<\theta_2\leq 1$, and the principal incurs a production cost of $0.5x^2$.

Consider now the binary signal whose signal realizations $\{s_1,s_2\}$ are drawn according to the conditional probability mass functions with $g(s_1|\theta_1)=\varepsilon$ and $g(s_1|\theta_2)=\varepsilon^2$ for the commonly known parameter $\varepsilon\in[0,1]$. The signal realizations generate the posteriors $g(\theta_1|s_1)=1/(1+\varepsilon)$ and $g(\theta_1|s_2)=1/(2+\varepsilon)$. We therefore obtain the inverse hazard rates $h(\theta_2|s_i)=0$, $h(\theta_1|s_1)=\varepsilon$, and $h(\theta_1|s_2)=$

$1+\varepsilon$. Moreover, suppose $\theta_1 > \frac{2}{3}\theta_2$. Under this condition, the quantities in the principal's optimal mechanism are strictly positive and given by $x(\theta_2|s_i) = \theta_2$, $x(\theta_1|s_1) = \theta_1 - \varepsilon\Delta\theta$, and $x(\theta_1|s_2) = \theta_1 - (1+\varepsilon)\Delta\theta$. Thus, we obtain an expected total surplus of

$$\begin{aligned}
&0.5(\varepsilon+\varepsilon^2)\left(\frac{1}{2(1+\varepsilon)}(\theta_1^2 - \varepsilon^2\Delta\theta^2) + \frac{\varepsilon}{2(2+\varepsilon)}\theta_2^2\right) \\
&\quad + 0.5(2-\varepsilon-\varepsilon^2)\left(\frac{1}{2(2+\varepsilon)}(\theta_1^2 - (1+\varepsilon)^2\Delta\theta^2) + \frac{1+\varepsilon}{2(1+\varepsilon)}\theta_2^2\right) \\
&= 0.25(\theta_1^2+\theta_2^2) - 0.25\Delta\theta^2\left(1+\varepsilon-\varepsilon^2\right).
\end{aligned} \tag{4}$$

Note that the term $1+\varepsilon-\varepsilon^2$ is strictly concave and takes its minima at $\varepsilon = 0$ and $\varepsilon = 1$. Thus, the uninformative signals maximize expected total surplus.

Suppose next that the set of production possibilities X is not a continuum but given by $X = \{0, \theta_1 - \Delta\theta, \theta_1, \theta_2\}$. I compare the uninformative signal with $\varepsilon = 0$ to an informative signal with $\varepsilon' \in (0,1)$. If ε' is sufficiently close to zero, the optimal quantities for the principal are given by $x(\theta_2|s_i) = \theta_2$, $x(\theta_1|s_1) = \theta_1$, and $x(\theta_1|s_2) = \theta_1 - \Delta\theta$. Thus, we obtain a Pareto improvement with respect to the uninformative signal under which the type θ_1 always receives the quantity $\theta_1 - \Delta\theta$. The discreteness of the set X implies that the slight change in the posterior belief for the signal realization s_2 from going from $\varepsilon = 0$ to $\varepsilon' > 0$ does not affect the quantities that the different types obtain after s_2. This contrasts with the case where X is a continuum. In this case, the increase in the likelihood of θ_2 after s_2 directly translates into a lower quantity for type θ_1 after s_2 which offsets the welfare gain for this type after s_1.

Data availability

No data was used for the research described in the article.

References

Aguirre, Inaki, Cowan, Simon, Vickers, John, 2010. Monopoly price discrimination and demand curvature. Am. Econ. Rev. 100 (4), 1601–1615.
Amir, Rabah, 2005. Supermodularity and complementarity in economics: an elementary survey. South. Econ. J. 71 (3), 636–660.
Athey, Susan, Levin, Jonathan, 2018. The value of information in monotone decision problems. Res. Econ. 72 (1), 101–116.
Baron, David P., Myerson, Roger B., 1982. Regulating a monopoly with unknown cost. Econometrica 50 (4), 911–930.
Bergemann, Dirk, Morris, Stephen, 2013. Robust predictions in games with incomplete information. Econometrica 81 (4), 1251–1308.
Bergemann, Dirk, Morris, Stephen, 2016. Bayes correlated equilibrium and the comparison of information structures in games. Theor. Econ. 11 (2), 487–522.
Bergemann, Dirk, Brooks, Benjamin, Morris, Stephen, 2015. The limits of price discrimination. Am. Econ. Rev. 105 (3), 921–957.
Bergemann, Dirk, Brooks, Benjamin, Morris, Stephen, 2017. First-price auctions with general information structures: implications for bidding and revenue. Econometrica 85 (1), 107–143.
Blackwell, David, 1951. Comparison of experiments. In: Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 93–102.
Blackwell, David, 1953. Equivalent comparisons of experiments. Ann. Math. Stat. 24 (2), 265–272.
Bose, Subir, Zhao, Jinhua, 2007. Optimal use of correlated information in mechanism design when full surplus extraction may be impossible. J. Econ. Theory 135 (1), 357–381.
Boyer, Marcel, Laffont, Jean-Jacques, 2003. Competition and the reform of incentive schemes in the regulated sector. J. Public Econ. 87 (9), 2369–2396.
Carroll, Gabriel, 2015. Robustness and linear contracts. Am. Econ. Rev. 105 (2), 536–563.
Cella, Michela, 2008. Informed principal with correlation. Games Econ. Behav. 64 (2), 433–456.
Chassang, Sylvain, 2013. Calibrated incentive contracts. Econometrica 81 (5), 1935–1971.
Chi, Chang Koo, 2014. The value of information and dispersion. PhD dissertation. The University of Wisconsin-Madison.
Crémer, Jacques, 1982. A simple proof of Blackwell's "comparison of experiments" theorem. J. Econ. Theory 27 (2), 439–443.
Demougin, Dominique M., Garvie, Devon A., 1991. Contractual design with correlated information under limited liability. Rand J. Econ. 22 (4), 477–489.
Dequiedt, Vianney, Martimort, David, 2015. Vertical contracting with informational opportunism. Am. Econ. Rev. 105 (7), 2141–2182.
Dewatripont, Mathias, Jewitt, Ian, Tirole, Jean, 1999. The economics of career concerns, part I: comparing information structures. Rev. Econ. Stud. 66 (1), 183–198.
Di Tillio, Alfredo, Ottaviani, Marco, Sørensen, Peter Norman, 2021. Strategic sample selection. Econometrica 89 (2), 911–953.
Diamond, Peter A., 1998. Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. Am. Econ. Rev. 88 (1), 83–95.
Drugov, Mikhail, 2010. Information and delay in an agency model. Rand J. Econ. 41 (3), 598–615.
Drugov, Mikhail, 2011. Ex Ante Information in Adverse Selection. Available at SSRN 1623823.
Fudenberg, Drew, Tirole, Jean, 1991. Game Theory. MIT Press, Cambridge, MA.
Ganuza, Juan-José, Penalva, Jose S., 2010. Signal orderings based on dispersion and the supply of private information in auctions. Econometrica 78 (3), 1007–1030.
Garrett, Daniel F., 2014. Robustness of simple menus of contracts in cost-based procurement. Games Econ. Behav. 87, 631–641.
Garrett, Daniel F., 2021. Payoff implications of incentive contracting. Theor. Econ. 16 (4), 1281–1312.
Gary-Bobo, Robert, Spiegel, Yossi, 2006. Optimal state-contingent regulation under limited liability. Rand J. Econ. 37 (2), 431–448.
Gjesdal, Frøystein, 1982. Information and incentives: the agency information problem. Rev. Econ. Stud. 49 (3), 373–390.
Grossman, Sanford J., Hart, Oliver D., 1983. An analysis of the principal-agent problem. Econometrica 51 (1), 7–46.
Guesnerie, Roger, 1981. On taxation and incentives: Further reflections on the limits to redistribution. Mimeo.
Haghpanah, Nima, Siegel, Ron, 2023. Pareto-improving segmentation of multiproduct markets. J. Polit. Econ. 131 (6), 1546–1575.
Hammond, Peter J., 1979. Straightforward individual incentive compatibility in large economies. Rev. Econ. Stud. 46 (2), 263–282.
Harris, Milton, Kriebel, Charles H., Raviv, Artur, 1982. Asymmetric information, incentives and intrafirm resource allocation. Manag. Sci. 28 (6), 604–620.
Herweg, Fabian, Müller, Daniel, 2014. Price discrimination in input markets: quantity discounts and private information. Econ. J. 124 (577), 776–804.
Holmström, Bengt, 1979. Moral hazard and observability. Bell J. Econ. 10 (1), 74–91.
Jewitt, Ian, 2007. Information Order in Decision and Agency Problems. Nuffield College. Mimeo.
Kim, Son Ku, 1995. Efficiency of an information system in an agency model. Econometrica 63 (1), 89–102.
Kim, Yonggyun, 2023. Comparing information in general monotone decision problems. J. Econ. Theory 211, 105679.
Laffont, Jean-Jacques, Tirole, Jean, 1986. Using cost observation to regulate firms. J. Polit. Econ. 94 (3), 614–641.

Lehmann, Erich Leo, 1988. Comparing location experiments. Ann. Stat. 16 (2), 521–533.
Maskin, Eric, Riley, John, 1984. Monopoly with incomplete information. Rand J. Econ. 15 (2), 171–196.
Mussa, Michael, Rosen, Sherwin, 1978. Monopoly and product quality. J. Econ. Theory 18 (2), 301–317.
Myerson, Roger B., 1981. Optimal auction design. Math. Oper. Res. 6 (1), 58–73.
Ottaviani, Marco, Prat, Andrea, 2001. The value of public information in monopoly. Econometrica 69 (6), 1673–1683.
Pigou, Arthur Cecil, 1920. The Economics of Welfare. Macmillan, London.
Quah, John K-H., Strulovici, Bruno, 2009. Comparative statics, informativeness, and the interval dominance order. Econometrica 77 (6), 1949–1992.
Riordan, Michael H., Sappington, David E.M., 1988. Optimal contracts with public ex post information. J. Econ. Theory 45 (1), 189–199.
Robinson, Joan, 1933. Economics of Imperfect Competition. Macmillan, London.
Saak, Alexander E., 2007. A note on the value of public information in monopoly. Econ. Theory 33 (2), 369–379.
Schmalensee, Richard, 1981. Output and welfare implications of monopolistic third-degree price discrimination. Am. Econ. Rev. 71 (1), 242–247.
Shaked, Moshe, Shanthikumar, J. George, 2007. Stochastic Orders. Springer.
Strausz, Roland, 2006. Deterministic versus stochastic mechanisms in principal–agent models. J. Econ. Theory 128 (1), 306–314.