Mechanism Design with Endogenous Information*

Anne-Katrin Roesler†

June 10, 2014

Abstract

In mechanism design problems with endogenous information, regularity properties of the distribution of posterior estimates (types) are essential for tractability. Important properties are a monotone hazard rate, increasing virtual valuations or costs. Difficulties arise since these properties are not preserved under mixtures, and regularity of the prior distribution may not translate to the distribution of posterior types. In this note, we identify sufficient conditions on the primitives of an information structure, which guarantee that the distribution of posterior types has a monotone hazard rate, increasing virtual valuations or costs. These characterization results make it possible to study mechanism design problems with endogenous information, without imposing regularity conditions on the interim stage or restricting attention to specific information structures. Applications to information acquisition and disclosure in optimal auctions, and to allocation problems without money are discussed.

Keywords: Information structures · Monotone hazard rate · Regularity · Distribution of posterior estimates · Mechanism design
JEL Classifications: C46, C60, D82

*I am grateful to Benny Moldovanu, Dirk Bergemann, and Larry Samuelson for their valuable comments and suggestions. I also thank Johannes Hörner, Daniel Krähmer, Jose Penalva and Bruno Strulovici for helpful comments and insightful discussions. Financial support from the DFG through Bonn Graduate School of Economics and from the German Academic Exchange Service (DAAD-Doktorandenstipendium) is gratefully acknowledged.

†Department of Economics, University of Bonn, 53113 Bonn, Germany, aroesler@uni-bonn.de.

1 Introduction

Consider a setting with endogenous information, in which the distribution of posterior types of agents emerges from the information acquisition or disclosure choices of the agents. In the process of Bayesian updating, mixtures over distributions are formed, an operation under which the increasing hazard rate property is not generally preserved. That is, for a prior distribution F with support $\mathcal{X} \subseteq \mathbb{R}$ and a family of distributions $\{G(\cdot|x)\}_{x\in\mathcal{X}}$ with support $S \subseteq \mathbb{R}$, even if all of these distributions have an increasing hazard rate this is not generally the case for the mixture distribution

$$G(s) = \int_{\mathcal{X}} G(s|x)\,\mathrm{d}F(x).$$

Consequently, even if the prior distribution of types has an increasing hazard rate, the distribution of posterior types induced by the endogenous choices of agents may not have this property.

In mechanism design settings, in which agents' information is endogenous, conditions that guarantee regularity of the distribution of posterior types are essential for tractability.[1] Without this assumption a circular effect could arise: small changes in the information level of agents could result in significant changes of the structure of the optimal mechanism, which would change the incentives to acquire or disclose information. This effect would render the model fragile, complicate the analysis tremendously, and make the model untractable. Under what conditions can we guarantee that all feasible choices of agents lead to regularity of the distribution of posterior types?

The main objective of this note is to identify sufficient conditions on the primitives of an information structure that guarantee that the distribution of posterior estimates has an increasing hazard rate, increasing virtual valuations or costs.[2] This characterization result is important for the emerging literature on mechanism design with endogenous information of agents. This literature dispenses with the common assumption that the distribution of types, and the private information held by agents, is exogenously given. It includes an infor-

[1]By contrast, in settings with exogenous information the role of regularity conditions is to simplify the analysis and avoid technicalities, specifically ironing-out procedures.

[2] It is a well-known problem in the economic literature that certain properties are not generally preserved under aggregation or mixtures. A prominent example is the single-crossing property introduced by Milgrom and Shannon (1994), which is not preserved under aggregation. Quah and Strulovici (2012) provide sufficient conditions that guarantee that the single crossing condition is preserved under aggregation. We provide a similar result: sufficient conditions for the increasing hazard rate property to be preserved under mixtures.

mation stage in the analysis, in which information is either acquired by market participants or disclosed to them.[3]

In our analysis, we focus on the standard setting for mechanism design problems, in which agents are risk-neutral and have quasi-linear preferences. In such a framework, all payoff-relevant information of agents that is necessary to characterize the optimal mechanism in the second stage is captured in the *posterior estimates (types)* of the agents. It is therefore not necessary to know the full posterior distribution, but it suffices to know its mean.

Our first result is an "impossibility result". We identify a class of signal structures for which the resulting distribution of posterior types will always have a decreasing hazard rate, irrespective of the prior distribution of types.[4] If a signal structure from this class is contained in the set of feasible signal structures that agents can choose from, it is impossible to guarantee that the induced distribution of posterior types has an increasing hazard rate for all feasible choices of agents.

Our second result is a "possibility result". We identify sufficient conditions on the signal structure that guarantee that certain regularity properties of the prior distribution – an increasing hazard rate, increasing virtual valuations or costs – translate to the distribution of posterior estimates.[5]

Straightforward applications of our results are the auction design problems analyzed in Shi (2012) and Ganuza and Penalva (2014). Shi (2012) studies optimal auctions with information acquisition by the bidders, whereas the focus of Ganuza and Penalva (2014) is on information disclosure in optimal auctions. The authors of these papers choose different approaches to circumvent the tractability problems that arise in their models. Shi (2012) imposes the regularity assumption directly on the distribution of posterior estimates, assuming that it has increasing virtual valuations. Ganuza and Penalva (2014) restrict attention to a specific information structure to make their model tractable. The results presented in this note make it possible to identify classes of information structures to which the results in Shi (2012) and Ganuza and Penalva (2014) apply. These applications are discussed in Section 4.

[3]Examples include Bergemann and Välimäki (2002) and Shi (2012) who study information acquisition, whereas the focus in Bergemann and Pesendorfer (2007), Esö and Szentes (2007), Ganuza and Penalva (2010), Li and Shi (2013) and Ganuza and Penalva (2014) is on information disclosure. Bergemann and Välimäki (2007) provide a good survey of the topic.

[4]This is the case for signal structures that are characterized by a family of conditional distributions that all have a decreasing hazard rate.

[5]Formally, signals must be characterized by a family of survival functions that is log-concave. This property is a generalization of the increasing hazard rate property to multivariate distributions.

To further illustrate how our results can be applied, we discuss information disclosure in allocation problems without monetary transfers.[6] We find that, by choosing an appropriate information technology, the designer can guarantee that the optimal mechanism is a full screening mechanism. This result is robust in the sense that the designer does not need to know the prior distribution of agents' types.

The rest of the note is organized as follows. In Section 2 we introduce the formal model of the informational environment. Section 3 contains our theoretical results, with the main results presented in Subsection 3.3. Applications are discussed in Section 4. We conclude with some further discussion and remarks in Section 5. All proofs are relegated to the appendix.

2 The Informational Setting

We consider the following model. There exists an unknown state, represented by a real-valued random variable X. The common, initial beliefs about the state are captured by an absolutely continuous prior distribution F with interval support $\mathcal{X} \subseteq \mathbb{R}$. We assume that X has finite expectation under F, $\mu = E(X) < \infty$.

A signal is characterized by a real-valued random variable S with typical realizations $s \in [\underline{s}, \overline{s}] \subseteq \mathbb{R}$, and a family of conditional distributions $\{G(s|x)\}_{x \in \mathcal{X}}$, where

$$G(s|x) := Pr(S \leq s | X = x)$$

is the probability to observe a signal $s' \leq s$ if the state is x.[7] We assume that for every $x \in \mathcal{X}$, $G(s|x)$, is absolutely continuous in s, that is, admits a density function $g(s|x) > 0$ almost everywhere.[8] Together with the prior distribution F, a signal induces a joint distribution on (X, S), a so-called *information structure*. We denote the marginal distribution of S by G.

Agents update their beliefs according to Bayes' rule. The posterior distribution of X conditional on observing s is $G(x|s)$, and the resulting conditional expectation is

$$\widehat{X}(s) = E\left[X | S = s\right] = \int_{\mathcal{X}} x \, \mathrm{d}G(x|s). \tag{1}$$

[6]For the case of exogenous private information of agents this problem has been studied for example in Condorelli (2012) and Chakravarty and Kaplan (2013).

[7]We allow for the supports of X and S to be the real line.

[8]This assumption implies that there is some noise in the signal. That is, upon observing a signal realization, agents cannot exclude any states. The set of states to which an agent attaches a positive probability is the same for all signal realizations. This assumption is sometimes called the "non-shifting support" assumption in the literature.

We call $\widehat{X}(s)$ the *posterior estimate*. Without loss of generality we can assume that $\widehat{X}(s)$ is increasing in s, which implies that an inverse function $\widehat{X}^{-1}$ exists.[9] For a given prior distribution F, every signal S results in a distribution of posterior estimates, represented by a random variable $\widehat{X} = E[X|S]$ with distribution function

$$H(\hat{x}) := G\left(\widehat{X}^{-1}(\hat{x})\right) = \int_{\mathcal{X}} G(\widehat{X}^{-1}(\hat{x})|x)\,\mathrm{d}F(x),$$

and quantile function $H^{-1}(p) = \inf\{\hat{x}|H(\hat{x}) \geq p\}$ for $p \in [0,1]$.

We assume that signals are monotone, that is, that high signal realizations are *more favorable* than low signal realizations in the sense of Milgrom (1981). This condition implies that it is more likely to observe a high signal realization s if the underlying state x is high, than if it is low.

Assumption 1 (Monotone Signals)**.** For all signal realizations $s, s' \in S$ with $s' > s$, signal realization s' is *more favorable* than s. That is, for every non-degenerate prior distribution F on X, if $s' > s$, then the posterior distribution $G(x|s')$ dominates $G(x|s)$ in terms of first-order stochastic dominance, $G(x|s') \geq_{FOSD} G(x|s)$.

If signal S is characterized by conditional densities $\{g(s|x)\}_{s\in\mathcal{S}}$, then Assumption 1 is equivalent to the monotone likelihood ratio property.[10]

Examples

Our model captures many information technologies. We now provide some examples that are frequently used in the literature.

Example 1 (Normal Experiments)**.** Suppose that the states are normally distributed $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, and signal S is given by $S = X + \varepsilon$ where ε is a normally distributed noise term, $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. It follows that signals are also normally distributed, $S \sim \mathcal{N}(\mu_X, \sigma_X^2 + \sigma_\varepsilon^2)$, and the posterior estimate after observing signal realization s is

$$\widehat{X}(s) = \frac{\sigma_\varepsilon^2}{\sigma_X^2 + \sigma_\varepsilon^2}\mu + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_\varepsilon^2}s.$$

The posterior estimates are linear in S and normally distributed. △

[9] For a formal justification see Shaked et al. (2012).

[10] Signal S has the (strict) *monotone likelihood ratio property* (MLRP), if for every $x > x'$, $\frac{g(s|x)}{g(s|x')}$ is strictly increasing in s.

Example 2 (Truth-or-noise technology)**.** Let the state space be X. The continuously differentiable distribution F with finite mean μ represents the prior beliefs. A *truth-or-noise technology* provides with some probability $\alpha \in [0,1]$ a perfectly informative signal $s = x$ and with probability $(1-\alpha)$ pure noise, independently drawn from prior distribution F. The receiver cannot distinguish which kind of signal he observes. For signal realization s, the posterior estimate is $\widehat{X}(s) = \alpha s + (1-\alpha)\mu$. △

Example 3. Suppose $X \sim U[0,1]$. If the state is x the resulting signal realizations are normally distributed with mean x and variance 1, that is, $G(s|x) \sim \mathcal{N}(x,1)$. The joint density which characterizes this information structure is

$$f(x,s) = g(s|x)f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{(s-x)^2}{2}} & \text{if } 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Upon observing signal realization s, the resulting posterior estimate is:

$$\widehat{X}(s) = s + \phi(0) \cdot (1-2s) = s \cdot (1 - 2\phi(0)) + \phi(0),$$

where $\phi(s) := \sqrt{\frac{2}{\pi}} \cdot \frac{\left[e^{-\frac{s^2}{s}} - e^{-\frac{(s-1)^2}{2}}\right]}{\operatorname{erf}\left(\frac{s}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{s-1}{\sqrt{2}}\right)}$, and erf is the error function.[11] Note that, as in the previous examples, the posterior estimate is linear in the signal realizations. △

3 Sufficient Conditions

In this section, we study the implications of properties of information structures for the distribution of posterior estimates, and identify sufficient conditions on the primitives of information structures for the distribution of posterior estimates to have a monotone hazard rate, increasing virtual valuations or costs.

Definition 1. The random variable X with distribution F and density f has an *increasing hazard rate*, if the *hazard rate function*

$$\lambda(x) = \frac{f(x)}{1-F(x)}$$

is increasing in x.
The random variable X has a *decreasing hazard rate*, if $\lambda(x)$ is decreasing in x.

[11] $\operatorname{erf}(s) = \frac{2}{\sqrt{\pi}} \int_0^s e^{-t^2}\, \mathrm{d}t$

An equivalent condition to X having an increasing (decreasing) hazard rate is that the *survival function* $\overline{F}(x) = 1 - F(x)$ is log-concave (log-convex).[12]

Remark 1. Interpreting the state x as time, the hazard rate $\lambda(x) = \frac{f(x)}{1-F(x)}$ has a natural interpretation as the *failure rate* of a component: It represents the probability of an instantaneous failure of a component conditional on the component still being intact at time x.

To establish our results we proceed in two steps. First, we identify sufficient conditions on the prior and signal distribution for the marginal distribution of signals to have an increasing hazard rate, respectively log-concave density (Lemma 1). We then show that these properties transfer to the distribution of posterior estimates (Proposition 1 and Proposition 2).

3.1 Induced properties of the marginal distribution of signals

A basic observation is, that the marginal distribution of signals is the mixture distribution over the conditional distributions characterizing the signal, with the prior being the mixing distribution

$$G(s) = \int_{\mathcal{X}} G(s|x)\, \mathrm{d}F(x).$$

It is a well-known result in statistics that the decreasing hazard rate property is preserved under mixtures.[13] For the increasing hazard rate property – the more important property for economics – the result is less clear-cut since the class of increasing hazard rate distribution is not closed under mixtures.

To develop some intuition about why the increasing hazard rate property is not necessarily preserved under mixtures, it is useful to think of the hazard rate function as representing the failure rate of a component (cf. Remark 1). A basic insight is that for mixtures of distributions early failures are likely to arise from distributions with high hazard rates. As Finkelstein and Cha (2013) put it "the weakest items are dying out first". More precisely, for a given prior distribution, consider the hazard rate of a mixture of a family of increasing (respectively decreasing) hazard rate distributions. For the mixture, early failures are more likely to arise from distributions with high hazard rates (at that time) whereas late failures are more likely to originate from low hazard rate distributions. This effect amplifies the features of decreasing hazard rate distributions but may offset the increasing hazard rate

[12] The natural definition of an increasing hazard rate for random variables without densities is, to say that X has an increasing hazard rate if the survival function is log-concave.

[13] See Barlow and Proschan (1981). In the appendix (Lemma 2) this result is formally stated.

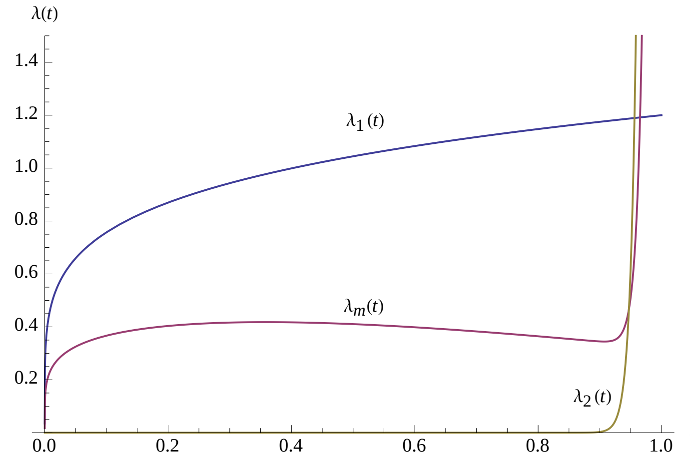

Figure 1: Hazard rate functions λ_1, λ_2 of Weibull distributions $\mathcal{W}(1.2, 1)$ and $\mathcal{W}(100, 1)$, with scale parameter 1 and shape parameters $k_1 = 1.2$ and $k_2 = 100$; and mixture hazard rate λ_m of their equal-weight mixture.

properties of the distributions when they are mixed. Consequently the increasing hazard rate property is not necessarily preserved under mixtures. Figure 1 illustrates an example of two distributions with increasing hazard rate whose mixture does not have this property.[14]

The following lemma identifies a set of sufficient conditions for the primitives of an information structure that guarantee that the marginal distribution of signals has an increasing hazard rate.[15]

Lemma 1. *Suppose the information structure* (X, S) *satisfies Assumption 1.*
If the family of survival functions $\{\overline{G}(s|x)\}_{x\in\mathcal{X}}$ *is log-concave in* (s, x)*, then if the prior distribution* F *has an increasing hazard rate, so has the marginal distribution of signals* G*.*
Moreover, if the family of densities $\{g(s|x)\}_{x\in\mathcal{X}}$ *is log-concave in* (s, x)*, then if the prior density* f *is log-concave, so is the marginal density of the signal* g*.*

3.2 Link to the distribution of posterior estimates

In order to obtain a general characterization result we still need to establish a relation between the marginal distribution of signals and the distribution of posterior estimates. For many of the information structures commonly used in the literature, among them the ones of Example 1 – 3, the posterior estimate is a positive linear transformation of the signal.

[14]For further examples, see Finkelstein and Cha (2013) and Gurland and Sethuraman (1994).

[15]This set is the least restrictive set of sufficient conditions we are aware of. The lemma is based on a theorem by Lynch (1999). For sufficient conditions for the case of a discrete state space see Block et al. (2003).

Assumption 2 (Linear Posterior Estimates)**.** Posterior estimates are a positive linear transformation of the signal:

$$\widehat{X} = aS + b,\ a, b \in \mathbb{R},\ a > 0.$$

For posterior estimates that do not satisfy this linearity condition, we impose the following smoothness condition.

Assumption 3 (Smoothness)**.** The distributions F and $\{G(\cdot|x)\}_{x\in\mathcal{X}}$ are twice continuously differentiable with strictly positive and bounded densities, $0 < f < \overline{f}$ and $0 < g(\cdot|x) < \overline{g}$ $\forall x \in \mathcal{X}$.

Many information structures satisfy both Assumption 2 and Assumption 3. We can now state our result, that for information structures satisfying at least one of these assumptions, the regularity properties of the marginal distribution of signals translate to the distribution of posterior estimates.

Proposition 1. *Suppose that the information structure (X, S) satisfies Assumption 1, either Assumption 2 or Assumption 3, and that the posterior estimate is a concave (convex) function of the signal. Then, if the marginal distribution of signals G has an increasing (decreasing) hazard rate, so has the distribution of posterior estimates H.*

Proposition 2. *For information structures satisfying Assumption 1 and Assumption 2, if the marginal density of signals g is log-concave, so is the density of posterior estimates, h.*

Remark. To avoid introducing new concepts and notation, the result of Proposition 2 is stated for log-concave densities. It should be noted, however, that the result extends to ρ-concave densities with ρ-concavity defined as in Caplin and Nalebuff (1991a,b) and Ewerhart (2013).

3.3 Main Results

Combining the results from Subsection 3.1 and Subsection 3.2, we can finally present our main result.

Theorem 1. *Suppose that the information structure (X, S) satisfies Assumption 1, and either Assumption 2 or Assumption 3.*

(i) If $\{G(s|x)\}_{x\in\mathcal{X}}$ is a family of decreasing hazard rate distributions and the posterior estimate is a convex function of the signal, then for any prior F, the distribution of posterior estimates H has a decreasing hazard rate.

(ii) If the family of survival functions $\{\overline{G}(s|x)\}_{x\in\mathcal{X}}$ is log-concave in (s,x) and the posterior estimate is a concave function of the signal, then if the prior distribution F has an increasing hazard rate, so does the distribution of posterior estimates H.

Remark. It should be noted that the conditions required to obtain the result in (ii) are significantly stronger than those in (i). Under the assumptions in the theorem, to guarantee that the distribution of posterior estimates has a decreasing hazard rate it suffices that for every $x \in \mathcal{X}$, the conditional distribution $G(\cdot|x)$ has a decreasing hazard rate. In particular, the result holds for any prior distribution. One can also think of this result as an "impossibility result": If signals are characterized by conditional distributions with a decreasing hazard rate, this makes it impossible that the increasing hazard rate property of the prior distribution translates to the distribution of posterior estimates.

By contrast, the possibility result for the increasing hazard rate in (ii), requires that the prior distribution has an increasing hazard rate and that the family of conditional survival distributions $\{\overline{G}(s|x)\}_{x\in\mathcal{X}}$ is log-concave in (s,x).[16]

We can also establish sufficient conditions for the virtual valuations and costs of the posterior estimates to be increasing.

Definition 2. A random variable X with distribution F and density f has *increasing virtual valuations* if

$$J_v(x) = x - \frac{1-F(x)}{f(x)},$$

is increasing in x.
It has *increasing virtual costs* if

$$J_c(x) = x + \frac{F(x)}{f(x)},$$

is increasing in x.

The following result is a direct corollary to Theorem 1.

Corollary 1. *Let (X,S) be an information structure that satisfies Assumption 1, and either Assumption 2 or Assumption 3.*
If the family of survival functions $\{\overline{G}(s|x)\}_{x\in\mathcal{X}}$ is log-concave in (s,x) and the posterior estimate is a concave function of the signal, then if the prior distribution F has an increasing hazard rate, the distribution of posterior estimates has increasing virtual valuations.

[16] A function $\psi : \mathbb{R}^2 \to \mathbb{R}$ is log-concave, if its domain dom ψ is convex and

$$\psi(\alpha x + (1-\alpha)y) \geq \psi(x)^{\alpha}\psi(y)^{1-\alpha} \quad \forall x, y \in \text{dom } \psi,\ \alpha \in (0,1)\,.$$

In order to guarantee that the distribution of posterior estimates has increasing virtual costs, slightly stronger conditions are required.

Theorem 2. *Let (X, S) be an information structure that satisfies Assumption 1 and Assumption 2.*
If the family of densities $\{g(s|x)\}_{x\in\mathcal{X}}$ is log-concave in (s, x), then if the prior density f is log-concave the distribution of posterior estimates has increasing virtual costs.[17]

4 Applications

4.1 Auctions

A straightforward starting point for the discussion of applications of our results, is to connect them to the existing research on auction design with endogenous information. Our results are effective in settings in which regularity of the posterior estimates does not arise from equilibrium considerations. This is for example the case in Shi (2012), who studies information acquisition in optimal auctions, as well as in Ganuza and Penalva (2014) who analyze information disclosure in optimal auctions.[18] In both settings, the implemented mechanism affects agents' incentives to acquire or disclose costly information and these informational effects have to be taken into account when designing the optimal mechanism. In both papers, the set of information technologies from which agents can choose is restricted, such that all information technologies in the feasible set can be compared in terms of their informational content.

4.1.1 Information Acquisition in Optimal Auctions

Shi (2012) characterizes an optimal, that is revenue maximizing, selling mechanism in a setting in which buyers do not know their private valuations ex-ante, but can acquire costly information prior to participating in the mechanism. The timing is as follows: 1. The seller

[17] The statement of the theorem can be strengthened. The conditions stated in the theorem imply that the generalized virtual valuation and cost functions $J_v(x) = x - \gamma\frac{1-F(x)}{f(x)}$ and $J_c(x) = x + \gamma\frac{F(x)}{f(x)}$ are increasing in x for every $\gamma > 0$.

[18] Bergemann and Pesendorfer (2007) study a setting in which the designer can choose what information to provide to bidders, and the selling mechanism. In this setting the non-decreasing virtual valuations property follows from equilibrium considerations: If providing information would result in non-increasing virtual valuations, the seller would rather not differentiate between buyers, but move the "ironing out" procedure to the information stage by not providing information to the agents. The result relies on the richness of the set of feasible information technologies that the designer may choose from. The information technologies in this set are not ordered in terms of informativeness and no predictions about the optimal precision-level of disclosed information can be made.

announces a mechanism (and suggests an information acquisition profile). 2. Bidders acquire costly information: they choose the precision level of the signal that they will obtain about their valuation of the object. 3. Based on their chosen precision levels of information, bidders obtain a (noisy) signal about their valuation and update their beliefs accordingly. 4. Bidders submit their bids, and the object is sold according to the mechanism previously announced by the seller.

In this environment, when choosing the optimal mechanism, the seller has to take into account that his choice of a mechanism will affect the incentives of bidders' to acquire information. For the symmetric case, in which all bidders acquire the same level of information, Shi (2012) shows that, if the distribution of posterior estimates is regular and the number of bidders is sufficiently large, the optimal mechanism is a standard auction with a reserve price. The optimal reserve price in the case with endogenous information acquisition is closer to the prior mean than the standard reserve price if the equilibrium information level were exogenously given.

Our results of Section 3 allow us to identify a class of information structures which satisfy the regularity condition on posterior estimates, necessary to establish the results of Shi (2012), without having to impose these regularity condition at the interim stage.[19] This class includes all information structures satisfying the conditions in Theorem 1 (*ii*) or Corollary 1. Examples include truth-or-noise technologies with increasing hazard rate prior distributions; and information structures with $S = X + \epsilon$, an increasing hazard rate prior distribution and noise with log-concave density.[20]

4.1.2 Information Disclosure in Optimal Auctions

Ganuza and Penalva (2014) study information disclosure in optimal auctions. In their setting, the seller chooses the selling mechanism, as well as the precision level of the information disclosed to bidders before the auction. The seller's choices are publicly observable, that is, known to all bidders. Prior to the auction, bidders observe a private, partially informative signal about their valuations for the object and update their beliefs accordingly before participating in the auction. The informational content of the signal is determined by the precision

[19] The result in Shi (2012) is based on further assumptions on the distribution of posterior estimates. A sufficient condition for these assumptions to hold is that, when switching from one signal to a more precise (and thus more costly) signal, the resulting distributions of posterior estimates are ordered in terms of the dispersive order. This is the case for information structures satisfying Assumption 2 (cf. Proposition 4 in Ganuza and Penalva (2010)).

[20] For classes of functions which have an increasing hazard rate or log-concave densities see Bagnoli and Bergstrom (2005).

level chosen by the seller.

In their analysis, Ganuza and Penalva (2014) assume that signals have the structure of a truth-or-noise technology (cf. Example 2). The authors state in their conclusion that "the model is standard (and general) in all dimensions but the choice of the set of available signals". This simplifying assumption that signals have the structure of a truth-or-noise-technology has the following convenient implications: 1. For truth-or-noise technologies, the regularity properties, like increasing virtual valuations of the prior distribution translate to the distribution of posterior estimates.[21] 2. Linearity of the information structure keeps the model tractable. 3. Information technologies are naturally ordered in terms of their informational content (precision) and it is straightforward to define a cost-function which captures the idea that information disclosure is costly.

If we allow for a larger set of information structures that satisfy Assumption 1 and Assumption 2 (linearity), the last two of the aspects mentioned above (2. and 3.) are preserved.[22] However, for these more general information structures the marginal distribution of signals is usually not the same as the prior distribution. Consequently, the increasing virtual valuation property will in general not translate from the prior distribution to the distribution of posterior estimates. In this case, our results of Section 3 can be applied to characterize sufficient conditions on the primitives of information structures for the distributions of posterior estimates to have the increasing virtual valuations property.

The results in Ganuza and Penalva (2014) generalize to the class of information structures satisfying Assumption 1, Assumption 2 and the conditions in Theorem 1 (*ii*) or Corollary 1.[23] The main insights are:

1. In an optimal auction, the auctioneer discloses more information than in a standard auction in which the object is always sold. Here, an optimal auctions is a standard auction with a reserve price that is optimal given the precision level of the information disclosed to bidders.

2. The level of information disclosed to bidders in an optimal auction is weakly increasing in the number of bidders.

[21]For truth-or-noise technologies, the marginal distribution of signals is the same as the prior distribution. Moreover, due to the linearity of the posterior estimates in signals, the regularity properties translate from the marginal distribution of signals to the distribution of posterior estimates.

[22]The class of information structures that satisfy Assumption 2, are naturally ordered in terms of *supermodular precision* (cf. Ganuza and Penalva (2010), Proposition 4).

[23]It is straightforward to replicate the proofs in Ganuza and Penalva (2014) for these more general information structures, using our results in Section 3 and the linearity of posterior estimates (Assumption 2). We refer the reader to the discussion in Ganuza and Penalva (2014).

The intuition behind these results is the following. In a standard auction (without reserve price), if a seller discloses information, he has to leave informational rents to the bidders. If information is costly, the auctioneer will therefore not reveal all information. A reserve price reduces the informational rents of bidders and thus increases the seller's incentives to disclose information.

4.2 Optimal Mechanisms without Money

Another interesting application of our results are allocation problems without monetary transfers as studied for example in Condorelli (2012) and Chakravarty and Kaplan (2013). These model consider the allocation of k indivisible heterogeneous objects to n agents when monetary transfers or charging personalized prices is infeasible or undesirable, but a benevolent designer can screen the agents.[24] Screening yields non-monetary costs which are wasted, that is, screening generates a deadweigh loss. More specifically, if the seller chooses to screen agents, he implements a mechanism that requires the agents to invest in some costly non-productive action (e.g. exerting effort, spend time in waiting lines, etc.) in order to signal their private types. Incentive compatibility requires that the non-monetary costs incurred by the agents correspond to Vickrey-payments. That is, the expected (wasteful) costs of an agent capture the externalities that he imposes on the other agents.

In this setting, Condorelli (2012) characterizes the optimal mechanism within the class of incentive compatible direct allocation mechanisms, that is, the mechanisms that maximizes ex-ante welfare. Condorelli shows that, if buyers' valuations have a decreasing hazard rate, a *full screening* mechanism is optimal whereas in any other case, due to the trade-off between a more efficient allocation and screening costs, only partial or no screening is optimal.

Our results of Section 3 can be applied to extend the model studied in Condorelli (2012) to a setting in which agents do not know their private valuations or tastes ex-ante and the seller can provide information through a noisy channel, for example by advertising a concert or sport event. We assume that the designer has to provide some information to make market participants aware of the availability of his products, but cannot provide perfectly informative private signals. For example, an event manager has to advertise a concert to attract interested customers but cannot perfectly control how interested parties perceive the information provided to them through the advertisement. Formally, this means that we

[24] Typical examples mentioned in the literature are the allocation of donor organs, or ticket sales for concerts or sport events. Waiting lines can serve as costly screening instruments.

exclude perfectly informative signals and pure noise from the set of feasible information technologies available to the designer. The setting studied in Condorelli (2012) is linear, hence the distribution of posterior estimates captures all relevant information to determine the optimal mechanism. Applying our insights from Section 3 yields the following result.

Corollary 2. *Suppose the seller implements a signal that is characterized by conditional distributions with decreasing hazard rate. Then for any prior distribution of agents' types, it is optimal for the seller to implement a full screening mechanism.*

We want to emphasize the following remarkable robustness feature of this result: As long as the designer can implement a signal characterized by decreasing hazard rate distributions, he knows that a full screening mechanism is optimal, irrespective of the prior type distribution of agents. To implement the optimal mechanism the designer therefore does not need to know the prior distribution.[25]

5 Discussion and Concluding Remarks

In this note we discussed properties of information structures and their implications for the distribution of posterior estimates. We specifically focused on identifying conditions such that the induced distribution of posterior estimates satisfies certain regularity properties that are commonly used in the mechanism design literature.

An important insight of the discussion is, that the increasing hazard rate property may not be preserved under mixtures of distribution functions, an operation which occurs during the updating process. For certain signal structures it is impossible that the distribution of posterior estimates has an increasing hazard rate. However, we identified sufficient conditions on the signal structure that guarantee that the increasing hazard rate property translates from the prior distribution to the distribution of posterior estimates.[26]

We used our results to identify classes of information environments to which the results on information acquisitions and disclosure in optimal auctions of Shi (2012) and Ganuza and

[25]The results of Corollary 2 extends to two-sided matching market models as discussed in Hoppe et al. (2009) and Roesler (2014). In these settings, if the designer implements an information technology characterized by decreasing hazard rate distributions, the welfare optimal mechanism is to screen agents and implement the positive assortative matching.

[26]As a corollary we obtain sufficient conditions for the distribution of posterior estimates to have increasing virtual valuations. However, these conditions are not tight and could probably be relaxed, using the insight from Ewerhart (2013) that $\left(-\frac{1}{2}\right)$-concavity is a tight sufficient condition for increasing virtual valuations, which is a weaker condition than an increasing hazard rate. The mathematical methods that we use to obtain our results do not extend to the case that would be needed to pursue this question systematically. We therefore second the statement of Hardy et al. (1952) that "the complications introduced by zero or negative values [are] hardly worth pursuing systematically".

Penalva (2014) apply. Moreover, we discussed information disclosure in allocation problems without money as for example studied in Condorelli (2012) or Chakravarty and Kaplan (2013). We showed that whenever signals are characterized by a family of decreasing hazard rate distributions, a full screening mechanisms is optimal.

We think that our results will be valuable beyond the applications discussed in this note, specifically for research on mechanism design problems with endogenous information. For such problems, the insights of Section 3 can be used to restrict attention to a set of feasible information structures for which the optimal mechanism is of a particular form. This allows to keep mechanism design problems with endogenous information tractable, an important first step to address new questions and develop new insights on this topic.

Appendix

Proof of Lemma 1. Suppose the information structure (X, S) satisfies the slightly stronger condition that f is log-concave. In this case, $\overline{G}(s, x) := \overline{G}(s|x)f(x)$ is log-concave, since the product of two log-concave functions is log-concave. The survival function of the marginal distribution G is given by:

$$\overline{G}(s) = \int_X \overline{G}(s|x)f(x)\,\mathrm{d}x.$$

By Prékopa's Theorem (1973), log-concavity is preserved by integration and it follows that $\overline{G}(s)$ is log-concave. Consequently, G has an increasing hazard rate (cf. Definition 1).

The same line of reasoning can be used to prove that log-concavity of f and $g(s|x)$ implies log-concavity of g.

For the proof of the general case, which only requires that F has an increasing hazard rate, observe that by Assumption 1, for every $s > s'$, $G(x|s) \geq_{FOSD} G(x|s')$. This implies, $G(s|x) \geq_{FOSD} G(s'|x)$ for all $x > x'$, which is equivalent to $\overline{G}(s|x) = 1 - G(s|x)$ being increasing in x for every $s \in S$. The result then follows by Theorem 2.1 in Lynch (1999). □

Proof of Proposition 1.

Case 1: Suppose the information structure (X, S) satisfies Assumption 1 and Assumption 2.[27] That is, suppose $\widehat{X} = aS + b$ with $a > 0$. This is equivalent to $S = \frac{\widehat{X}-b}{a}$. Given monotonicity of signals, for every $\hat{x} \in \widehat{\mathcal{X}}$, $\widehat{X}(s) \leq \hat{x} \Leftrightarrow s \leq \frac{\hat{x}-b}{a}$. This implies $H(\hat{x}) = G\left(\frac{\hat{x}-b}{a}\right)$.

Let $\eta(\hat{x}) := \frac{\hat{x}-b}{a}$. For $a > 0$, $\eta(\hat{x})$ is increasing in $\hat{x}$. Moreover,

$$h(\hat{x}) = \frac{\mathrm{d}G}{\mathrm{d}\eta}\frac{\mathrm{d}\eta}{\mathrm{d}\hat{x}} = \frac{1}{a} \cdot g\left(\eta(\hat{x})\right). \tag{2}$$

[27] In this case, the posterior estimate is a concave and convex function of the signal.

It follows that

$$\lambda_H(\hat{x}) = \frac{1}{a} \cdot \frac{g(\eta(\hat{x}))}{1 - G(\eta(\hat{x}))}.$$

Given that $a > 0$ and $\eta(\hat{x})$ is increasing in $\hat{x}$, it follows that, if G has an increasing (decreasing) hazard rate then $\lambda_H(\hat{x})$ is increasing (decreasing) in $\hat{x}$ which implies that H (resp. $\widehat{X}$) has an increasing (decreasing) hazard rate.

Case 2: Suppose the information structures (X, S) satisfies Assumption 1 and Assumption 3. These conditions imply that $\widehat{X}(s)$ is continuously differentiable and strictly increasing in s. By the inverse function theorem, $\hat{X}$ is invertible. That is, there exists a twice continuously differentiable function $\eta := \widehat{X}^{-1}$, and the first and second derivative of η are given by

$$\eta'(\hat{x}) = \frac{1}{\widehat{X}'(\eta(\hat{x}))} \quad \text{and} \quad \eta''(\hat{x}) = -\frac{\widehat{X}''(\eta(\hat{x}))}{\left(\widehat{X}'(\eta(\hat{x}))\right)^3}.$$

By Assumption 1, $\eta(\hat{x})$ is strictly increasing in $\hat{x}$. This implies that $\eta'(\hat{x}) > 0$ for every $\hat{x}$. Moreover, if $\widehat{X}$ is concave, then $\eta''(\hat{x}) > 0$, that is, η is convex. Similarly, if $\widehat{X}$ is convex, then $\eta''(\hat{x}) < 0$ and η is concave.

For every $\hat{x}$, $\eta(\hat{x})$ determines the signal realization s that results in the conditional expectation $\hat{x}$. With these specifications, $S = \eta(\widehat{X})$ and $H(\hat{x}) = G(\eta(\hat{x}))$. Since G and η are both continuously differentiable so is H, and it follows that

$$h(\hat{x}) = \frac{\mathrm{d}G}{\mathrm{d}\eta}\frac{\mathrm{d}\eta}{\mathrm{d}\hat{x}} = g(\eta(\hat{x})) \cdot \eta'(\hat{x}).$$

It follows that the hazard rate function of $\widehat{X}$ is given by

$$\lambda_H(\hat{x}) = \eta'(\hat{x})\frac{g(\eta(\hat{x}))}{1 - G(\eta(\hat{x}))}.$$

Its derivative is

$$\lambda'_H(\hat{x}) = \eta''(\hat{x})\psi(\hat{x}) + \eta'(\hat{x}) \cdot \psi'(\hat{x}),$$

with $\psi(\hat{x}) := \frac{g(\eta(\hat{x}))}{1-G(\eta(\hat{x}))}$.

If G has an increasing (decreasing) hazard rate, then $\psi(\hat{x}) = \frac{g(\eta(\hat{x}))}{1-G(\eta(\hat{x}))}$ is increasing (decreasing). Given that $\eta'(\hat{x}) > 0$ for every $\hat{x}$, it follows that the second summand of $\lambda'_H(\hat{x})$ is increasing (decreasing) in $\hat{x}$. Moreover, if $\widehat{X}$ is concave (convex) then $\eta''(\hat{x}) > 0$ ($\eta''(\hat{x}) < 0$), which implies that the first summand is increasing (decreasing).

It follows that the distribution of posterior estimates H has an increasing hazard rate, if the posterior estimate is a concave function of the signal and the distribution of signals has an increasing hazard rate. Similarly, the distribution of posterior estimates has a decreasing hazard rate, if the distribution of signals has a decreasing hazard rate and the posterior estimate is a convex function of the signal. □

Proof of Theorem 1.

(i) As mentioned in Subsection 3.1, it is a well-known results that the decreasing hazard rate property is preserved under mixtures. Formally

Lemma 2 (Barlow and Proschan (1981))**.** *Consider a family of distributions* $\{G_\theta(t)\}_{\theta\in\Theta}$ *that all have a decreasing hazard rate. Then, for any mixing distribution* F*, the mixture distribution*

$$G(t) = \int_\Theta G_\theta(t)\, dF(\theta)$$

has a decreasing hazard rate.

It is straightforward to apply this result to information structures: Let $\{G(s|x)\}_{x\in\mathcal{X}}$ be the family of distributions and the prior F be the mixing distribution. Then, if for every $x \in \mathcal{X}$, the distribution $G(s|x)$ has a decreasing hazard rate, Lemma 2 implies that the marginal distribution of signals $G(s)$ has a decreasing hazard rate. By Proposition 1 it follows that the distribution of posterior estimates, H, has a decreasing hazard rate.

(ii) The result follows directly by combining Lemma 1 and Proposition 1. Given the assumptions, Lemma 1 implies that the marginal distribution of signals G has an increasing hazard rate; by Proposition 1 the distribution of posterior estimates H has the same property. □

Proof of Corollary 1. The result follows directly from Theorem 1 and the fact, that an increasing hazard rate implies increasing virtual valuations. □

Proof of Theorem 2. Under the assumptions of the theorem, Lemma 1 implies that the marginal density function of signals g is log-concave. Applying Proposition 1 yields that the distribution of posterior estimates h is log-concave. It is shown in Ewerhart (2013) that log-concavity is equivalent to $\rho-$concavity for $\rho = -\frac{1}{2}$, and that this is a sufficient condition for the virtual cost function $J_c(\hat{x})$ to be increasing in $\hat{x}$. □

Proof of Corollary 2. Suppose the seller implements a signal that is characterized by conditional distributions with decreasing hazard rate, that is, for every $x \in \mathcal{X}$, the distribution

$G(s|x)$ has a decreasing hazard rate. Then, by Theorem 1 (i), the distribution of posterior estimates, H, has a decreasing hazard rate. By Theorem 1 and Corollary 3 in Condorelli (2012) it follows that the optimal mechanism is a full screening mechanism. □

References

Bagnoli, Mark and Ted Bergstrom (2005), "Log-Concave Probability and its Applications." *Economic Theory*, 26, 445–469.

Barlow, Richard E. and Frank Proschan (1981), *Statistical Theory of Reliability and Life Testing. Probability Models.* Holt, Rinehart and Winston Inc., New York.

Bergemann, Dirk and Martin Pesendorfer (2007), "Information Structures in Optimal Auctions." *Journal of Economic Theory*, 137, 580–609.

Bergemann, Dirk and Juuso Välimäki (2002), "Information Acquisition and Efficient Mechanism Design." *Econometrica*, 70, 1007–1033.

Bergemann, Dirk and Juuso Välimäki (2007), "Information in Mechanism Design." In *Proceedings of the 9th World Congress of the Econometric Society*, 186–221.

Block, Henry W., Thomas H. Savits, and Eshetu T. Wondmagegnehu (2003), "Mixtures of Distributions with Increasing Linear Failure Rates." *Jounal of Applied Probability*, 40, 485–504.

Caplin, Andrew and Barry Nalebuff (1991a), "Aggregation and Social Choice: A Mean Voter Theorem." *Econometrica*, 59, 1–23.

Caplin, Andrew and Barry Nalebuff (1991b), "Aggregation and Imperfect Competition: On the Existence of Equilibrium." *Econometrica*, 59, 25–59.

Chakravarty, Surajeet and Todd R. Kaplan (2013), "Optimal Allocation without Transfer Payments." *Games and Economic Behavior*, 77, 1–20.

Condorelli, Daniele (2012), "What money can't buy: Efficient Mechanism Design with Costly Signals." *Games and Economic Behavior*, 75, 613–624.

Esö, Peter and Balazs Szentes (2007), "Optimal Information Disclosure in Auctions and the Handicap Auction." *Review of Economic Studies*, 74, 705–731.

Ewerhart, Christian (2013), "Regular Type Distributions in Mechanism Design and ρ-concavity." *Economic Theory*, 53, 591–603.

Finkelstein, Maxim and Ji Hwan Cha (2013), *Stochastic Modeling for Reliability: Shocks, Burn-in and Heterogeneous populations.* Springer Series in Reliability Engineering, Springer.

Ganuza, Juan-Jose and Jose S. Penalva (2010), "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions." *Econometrica*, 78, 1007–1030.

Ganuza, Juan-Jose and Jose S. Penalva (2014), "Information Disclosure in Optimal Auctions." Working Paper 14-02, Business Economics Series, Universidad Carlos III de Madrid.

Gurland, John and Jayaram Sethuraman (1994), "Reversal of Increasing Failure Rates When Pooling Failure Data." *Technometrics*, 36, 416–418.

Hardy, Godfrey H., John E. Littlewood, and George Polya (1952), *Inequalities*. Cambridge University Press.

Hoppe, Heidrun, Benny Moldovanu, and Aner Sela (2009), "The Theory of Assortative Matching Based on Costly Signals." *The Review of Economic Studies*, 76, 253–281.

Li, Hao and Xianwen Shi (2013), "Discriminatory Information Disclosure." Working Paper, University of Toronto.

Lynch, James D. (1999), "On Conditions for Mixtures of Increasing Failure Rate Distributions to have an Increasing Failure Rate." *Probability in the Engineering and Informational Sciences*, 13, 33–36.

Milgrom, Paul R. (1981), "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics*, 12, 380–391.

Milgrom, Paul R. and Chris Shannon (1994), "Monotone Comparative Statics." *Econometrica*, 62, 157–180.

Prékopa, Andras (1973), "On Logarithmic Concave Measures and Functions." *Acta Scientiarum Mathematicarum*, 34, 335–343.

Quah, John K.-H. and Bruno Strulovici (2012), "Aggregating the single crossing property." *Econometrica*, 80, 2333–2348.

Roesler, Anne-Katrin (2014), "Information Disclosure in Markets: Auctions, Contests, and Matching Markets." Working Paper.

Shaked, Moshe, Miguel A. Sordo, and Alfonso Suarez-Llorens (2012), "Global Dependence Stochastic Orders." *Methodology and Computing in Applied Probability*, 14, 617–648.

Shi, Xianwen (2012), "Optimal Auctions with Information Acquisition." *Games and Economic Behavior*, 74, 666–686.