# Extending Blackwell and Lehmann: The Monotone Quasi-Garbling Order

Yonggyun Kim*

Duke University

yonggyun.kim@duke.edu

September 2, 2018

**Abstract**

How should one evaluate information in a monotone decision problem where a higher action is optimal for a higher signal realization? As a criterion for comparing information structures in such environments, I develop a condition called *monotone quasi-garbling* meaning that an information structure is obtained by adding *reversely monotone noise* to another. This new criterion generally permits more comparisons than the garbling condition by Blackwell [1951, 1953] and the effectiveness condition by Lehmann [1988]. For a general class of monotone decision problems, it is shown that monotone quasi-garbling is a sufficient condition for decision makers to get a higher ex-ante expected payoff. To illustrate, I apply the result to an optimal insurance problem and to a nonlinear monopoly pricing problem.

*JEL Classification:* C44, C60, D81, D86

*Keywords:* Information, monotonicity, Blackwell's condition, Lehmnann's condition, optimal insurance, nonlinear pricing.

*I am grateful to Pino Lopomo, Todd Sarver, and Curtis Taylor for excellent comments and suggestions. I would also like to thank Yuhta Ishii, Jinwoo Kim, Son Ku Kim, Peter Norman Sørenson, Feifan Zhang, and seminar participants at Duke University and Seoul National University for helpful comments. Any remaining errors are my responsibility.

# 1 Introduction

Consider a pair of information structures that provide signals about uncertain states. When can we say that one is superior to the other? This fundamental question has numerous economic applications such as investment, monopoly pricing, and auctions. A common feature in settings such as these is that the decision maker would like to take a higher action when a higher signal is realized, that is, the decision problem is often *monotone*.

The classical way of comparing information structures is to use the *garbling condition* developed by Blackwell [1951, 1953]. By this criterion an information structure ($g$) is worse than another ($f$) if $g$ can be obtained from $f$ by adding some noise–in other words, $g$ is a garbling of $f$. Intuitively, the added noise reduces the value of information. Indeed, Blackwell's condition implies that for *any* preferences that satisfy the von Neumann-Morgenstern axioms, the expected payoff under $f$ is higher than under $g$. This is a powerful result because once a pair of information structures is ranked by Blackwell's condition, the rank is preserved in almost every decision problem. At the same time, it is hard to satisfy Blackwell's condition, so this partial ordering of information structures has limited applicability.

If we restrict attention to monotone decision problems, it is sometimes possible to compare pairs of information structures that are unrankable by Blackwell's condition. One well-known criteria using this approach is *the effectiveness condition* by Lehmann [1988]. Lehmann refines Blackwell's condition by assuming that (i) the decision maker has monotone preferences; and (ii) both information structures have the monotone likelihood ratio property (MLRP).[1]

In this paper, I introduce a novel criterion called *monotone quasi-garbling* to overcome a potential shortcoming of both Blackwell's and Lehmann's conditions in the following scenario. Consider three information structures $f$, $g$ and $h$, and assume that $f$ is more Blackwell sufficient than $g$ and $g$ is more Lehmann effective than $h$. In addition, assume that $g$ and $h$ satisfy the MLRP, but $f$ does not. In this scenario, we can ascertain that $f$ is more informative than $h$ for monotone decision problems because $f$ is more informative than $g$ for all decision problems by Blackwell's result and $g$ is more informative than $h$ for monotone decision problems by Lehmann's result. However, when the decision maker does not know the existence of $g$ and directly

[1] For the precise definition of the monotone likelihood ratio property, see Definition 2.4.

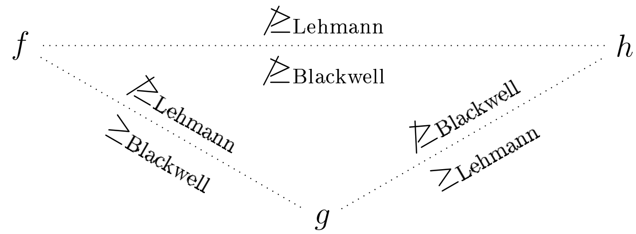


Figure 1: An example illustrating a shortcoming of Blackwell and Lehmann

compares $f$ and $h$, it is possible that neither Blackwell's nor Lehmann's condition shows that $f$ is more informative than $h$ as in Figure 1. A concrete example of such a case is provided in the next subsection. The example shows that there exists a pair of information structures such that one guarantees higher expected payoffs than another for any monotone decision problem, but they are not rankable by the existing criteria.

The key factor that generates this gap is the absence of the MLRP of $f$. Recall that Lehmann's effectiveness condition refines Blackwell's garbling condition when both information structures have the MLRP. Hence, without the MLRP of $f$, even though $f$ is more Blackwell sufficient than $g$, $f$ may not be more Lehmann effective than $g$. This observation calls for a criterion such that (i) the new criterion is weaker than both Blackwell's and Lehmann's conditions; (ii) it works for monotone decision problems without imposing the MLRP on the supposedly higher information structure. The main result of the paper is that monotone quasi-garbling satisfies both requirements. It is shown that monotone quasi-garbling is weaker than Blackwell's garbling condition and Lehmann's effectiveness condition (Proposition 2.1 and Proposition 2.3), but with the MLRP of the higher information structure, monotone quasi-garbling is equivalent to Lehmann's effectiveness condition (Proposition 2.4). It is also shown that $f$ and $h$ in Figure 1 are comparable by the monotone quasi-garbling criterion (Proposition 2.5). Moreover, without imposing the MLRP on the higher information structure, it is shown that monotone quasi-garbling implies more informativeness for monotone decision problems (Theorem 1). In addition, for some special cases, monotone quasi-garbling is also a necessary condition for more informativeness for monotone decision problems (Theorem 2, 3, 4).

The monotone quasi-garbling criterion also provides a fresh viewpoint for comparing information structures for monotone decision problems by introducing the concept of *reversely monotone noise*. Recall that the key idea of Blackwell's condition is that an information structure is deteriorated by adding noise. One of the underlying assumptions in Blackwell's condition is that the added noise is independent of the state. In the monotone quasi-garbling condition, I relax this assumption as follows: (i) I allow noise to be state dependent; (ii) but I restrict noise to be reversely monotone in the sense that higher noise corresponds for a lower state and lower noise corresponds for a higher state. Intuitively, the order of states matters in monotone decision problems, thus, an information structure can be degraded by adding higher noise to a lower state and lower noise to a higher state.

Another contribution of the paper is to provide a general framework for monotone decision problems that can be applied even in cases of multidimensional decisions. In the previous literature on monotone decision problems, it is usually assumed that the decision maker must choose a single dimensional alternative so that the set of alternatives has a generic order. On top of that, the MLRP is assumed, making a higher alternative optimal for the decision maker when a higher signal arrives. Although previous studies have also explored monotone decision problems, the assumption of single dimensionality restricts the class of problems that can be dealt with. For example, consider a monopolistic seller who faces an uncertain type of buyer. Before the monopolist makes a pricing offer, he can receive a report about the buyer (regard this as a signal) from an analyst (regard this as an information structure). Then, the seller updates his belief about the buyer's type based on the report and offers a nonlinear pricing schedule as explored by Mussa and Rosen [1978] and Maskin and Riley [1984].[2] In this example, a menu of prices and quantities can be interpreted as a multidimensional decision. To develop a theory that can also be applied for multidimensional sets of alternatives, I introduce a *well-ordered condition* (corresponding to the generic order of the single dimensional decision space) and a *monotonicity condition* (corresponding to the MLRP of the information structure).

In the rest of this section, I give an example illustrating the gap not covered by Blackwell and Lehmann and then review related literature. In Section 2, the monotone quasi-garbling concept is introduced and compared to the conditions of Blackwell and Lehmann. In Section 3, a general class of monotone decision problems is described and it is shown that monotone quasi-garbling is a sufficient condition for more informativeness for monotone decision problems and is also a necessary

[2] This application will be studied in detail in Subsection 4.3.

| | Fred ($f$) | | | | George ($g$) | | Hannah ($h$) | |
|---|---|---|---|---|---|---|---|---|
| States | LL | HL | LH | HH | L | H | L | H |
| $\omega_1$ | 1/2 | 1/2 | 0 | 0 | 1 | 0 | 5/6 | 1/6 |
| $\omega_2$ | 1/2 | 0 | 1/2 | 0 | 1/2 | 1/2 | 2/3 | 1/3 |
| $\omega_3$ | 0 | 1/2 | 0 | 1/2 | 1/2 | 1/2 | 1/2 | 1/2 |
| $\omega_4$ | 0 | 0 | 1/2 | 1/2 | 0 | 1 | 1/6 | 5/6 |

Figure 2: An example of information structures for three analysts

condition in some cases. In Section 4, the result is applied to the optimal insurance and the nonlinear monopoly pricing problems. Section 5 concludes.

## 1.1 An Illustrative Example

Consider three independent analysts Fred ($f$), George ($g$), and Hannah ($h$), whose information structures are given as in Figure 2. I assume that Fred can observe George's signal so that Fred's information structure in the figure is given by the joint information of Fred's own information and George's. To be more specific, Fred's own information structure is given as follows: a half for each signal for $\omega_1$ and $\omega_4$, L for sure for $\omega_2$, H for sure for $\omega_3$. Since Fred's and George's information are independent, the joint information is given as in the figure.

A simple application of Blackwell's condition is that when two analysts are independent, receiving information from both analysts is better than receiving information from only one of them, thus, $f$ is more Blackwell sufficient than $g$. Moreover, $g$ is more Lehmann effective than $h$. Also note that $g$ and $h$ satisfy the MLRP, but $f$ does not. In Appendix A.2, I show the following observations on these three information structures:

1. $f$ is more Blackwell sufficient than $g$, but $f$ is not more Lehmann effective than $g$;

2. $g$ is more Lehmann effective than $h$, but $g$ is not more Blackwell sufficient than $h$;

3. $f$ is neither more Blackwell sufficient nor Lehmann effective than $h$;

4. $h$ is a monotone quasi-garbling of $f$, in other words, $f$ is better than $h$ by the monotone quasi-garbling criterion introduced in Definition 2.2 below.

Note that this example satisfies the relations described in Figure 1 by the above observations. By the first observation and Blackwell's theorem, $f$ is more informative than $g$ for all decision problems, and by the second observation and Lehmann's theorem, $g$ is more informative than $h$ for monotone decision problems. Therefore, by transitivity, $f$ should be more informative than $h$ for monotone decision problems, but by the third observation, neither Blackwell's nor Lehmann's theorem can be applied for the direct comparison of $f$ and $h$.

On the other hand, the fourth observation means that $f$ and $h$ are comparable by the monotone quasi-garbling criterion. Then, the main theorem of the paper (Theorem 1) can be applied in comparing $f$ and $h$ and it shows that $f$ is more informative than $h$ for monotone decision problems.

## 1.2 Related Literature

Since Blackwell [1951, 1953] introduced a criterion to compare information structures in general decision problems, subsequent studies have refined this criterion by restricting it to monotone decision problems. In a seminal paper, Lehmann [1988] mentions location experiments to point out the limit of Blackwell's condition. Then, he restricts decision problems to have the MLRP for information structures and monotone utility introduced by Karlin and Rubin [1956] and establishes a theorem that his effectiveness condition is a necessary and sufficient condition for informativeness. Persico [2000] utilizes this effectiveness condition for decision problems with single crossing utility.[3] Quah and Strulovici [2009] introduce the interval dominance order property which is weaker than the single crossing property and shows that Lehmann's condition can also be utilized as a criterion for decision problems with this property. These studies are well summarized by a unified framework provided by Chi [2015]. He establishes an equivalence of effectiveness, informativeness, and posterior dispersion under decision problems with supermodular, single crossing and interval dominance orders.

In this paper, I establish a criterion for monotone decision problems without imposing the MLRP on an information structure. This possibility is explored by Quah

[3] Jewitt [2007] compares Karlin Rubin monotonicity and the single crossing property.

and Strulovici [2009]. In Theorem 3 of their paper, when the decision maker's preference has the interval dominance order, they show that Lehmann's effectiveness condition implies informativeness by only assuming the MLRP of the supposedly lower information structure. The main result of this paper extends their result in that (i) more comparisons are possible (monotone quasi-garbling is weaker than the effectiveness condition); (ii) it can be applied to a broader class of monotone decision problems (monotone decision problems that this paper deals with include decision problems with the interval dominance order).

A property of the monotone quasi-garbling criterion is that the ordering is independent of prior beliefs. It is also true for conditions of Blackwell and Lehmann. On the other hand, several recent studies exploit prior beliefs to refine Lehmann's condition. Athey and Levin [2017] restrict utility functions to satisfy certain conditions such as supermodularity and fix a prior belief, then introduce a criterion called monotone information order. Ganuza and Penalva [2010] apply integral and supermodular orders, which are defined on probability measures on expectations of states, to auction problems. Note that the expectation of states largely depends on the prior belief. Lastly, Cabrales et al. [2013] restricts attention to an investment decision problem and shows that an entropy ordering, which is prior dependent, gives a complete informative ordering. Since these orderings utilize prior beliefs as additional sources for information comparisons, they perform better than prior independent orderings if the prior is known. Nevertheless, it is still important to establish prior independent orderings because it is essential for applications with unknown or multiple prior beliefs.

The comparison of information structures has been applied in numerous economic situations: investment decisions (Cabrales et al. [2013]), auctions (Persico [2000], Ganuza and Penalva [2010]), matching markets (Roesler [2015]), principal agent models with moral hazard (Kim [1995]), competitive markets with adverse selection (Levin [2001]), and monopoly pricing (Athey and Levin [2017], Ottaviani and Prat [2001]). I apply the main result of this paper to an optimal insurance problem, which is a variant of the investment decision problem by Cabrales et al. [2013], and to a nonlinear monopoly pricing problem.

## 2 Information Ranking Criteria

Let $\Omega \subset \mathbb{R}$ be the set of states of nature and assume that $\Omega$ is compact. Denote $\omega \in \Omega$ as a generic state. An information structure is composed of a compact set of signals $X \subset \mathbb{R}$ and a collection of signal distributions $\{f(\cdot|\omega)\}_{\omega\in\Omega}$ with $f(\cdot|\omega) \in \Delta(X)$ where $\Delta(X) \equiv \{p : X \to [0,1] \mid \int_X p(x)dx = 1\}$. Let the cumulative distribution function of $f(\cdot|\omega)$ be $F(\cdot|\omega)$. Simply denote this information structure as $f$. For another information structure $g$, let the set of signals be $Y \subset \mathbb{R}$.

As a criterion for comparing information structures, I introduce a novel condition called monotone quasi-garbling and explore how it is related to Blackwell's garbling condition and Lehmann's effectiveness condition. I begin by reviewing Blackwell's garbling notion briefly.

**Definition 2.1** (Garbling by Blackwell [1951, 1953])**.** An information structure $g$ is said to be a *garbling* of $f$ or $f$ is *more Blackwell sufficient* than $g$, denoted $f \geq_B g$, if there exists a function $\gamma : X \to \Delta(Y)$ such that

$$g(y|\omega) = \int_{x\in X} \gamma(y|x)dF(x|\omega).$$

As an interpretation of this relation, a signal $Y$ with pdf $g$ is constructed by adding noise $\gamma$ to a signal $X$ with pdf $f$ so that $f$ is better information than $g$. In the garbling relation, the noise represented by $\gamma$ is only dependent on a signal $x$, not on a state $\omega$. Therefore, the noise in Blackwell's condition can be considered as *unilateral* noise. In the monotone quasi-garbling criterion, I allow the noise to be state dependent but restrict the noise to have some monotone relation.

**Definition 2.2** (Monotone Quasi-garbling)**.** An information structure $g$ is said to be a *monotone quasi-garbling* of $f$, denoted $f \geq_{MQG} g$, if there exists a function $\gamma : X \times \Omega \to \Delta(Y)$ such that

1. $g(y|\omega) = \int_{x\in X} \gamma(y|x,\omega)dF(x|\omega)$,

2. $\gamma(\cdot|x,\omega) \geq_{FOSD} \gamma(\cdot|x,\omega')$ for all $\omega' > \omega$ and $x \in X$.[4]

[4]When $H$ and $H'$ are cdfs of $h$ and $h'$, respectively, then $h(\cdot) \geq_{FOSD} h'(\cdot)$ iff $H(z) \leq H'(z)$ for all $z$.

The first condition is the same as the garbling notion except that the noise may depend on $\omega$. If there were no restriction on $\gamma$, any information structure $g$ could be generated by adding state contingent noise to $f$ and it would be a meaningless criterion. The second condition restricts state contingent noise to be *reversely FOSD-ordered.* This means that the noise would return a higher $y$ for a lower state and lower $y$ for a higher state, thus, it can be interpreted to be *reversely monotone.* Therefore, we can ascertain that the noise would lower the quality of information so that $f$ is still better information than $g$.

As a simple observation, note that if $g$ is a garbling of $f$, then $g$ is a monotone quasi-garbling of $f$ as well because a probability distribution is first order stochastic dominant over itself. However, by the example of information structures for Fred and Hannah in Subsection 1.1, we can observe that monotone quasi-garbling does not necessarily imply garbling. Therefore, monotone quasi-garbling is a weaker condition than Blackwell's garbling condition and this observation is summarized in the following proposition.

**Proposition 2.1.** *For any pair of information structures $(f, g)$, $f \geq_B g$ implies $f \geq_{MQG} g$, but $f \geq_{MQG} g$ does not imply $f \geq_B g$.*

Next, I review the definition of Lehmann's effectiveness condition and explore how it is related to Blackwell's condition and monotone quasi-garbling.

**Definition 2.3** (Effectiveness by Lehmann [1988]). An information structure $f$ is said to be *more Lehmann effective* than $g$, denoted $f \geq_L g$, if, for all $\omega < \omega'$, and all signals $y \in Y$,

$$F^{-1}(G(y|\omega)|\omega) \leq F^{-1}(G(y|\omega')|\omega'),$$ [5]

equivalently,

$$T(\omega; y) = F^{-1}(G(y|\omega)|\omega) \text{ is nondecreasing in } \omega.$$ [6]

An assumption that is usually paired with Lehmann's effectiveness condition is the monotone likelihood ratio property.

[5] When the type space is discrete, Lehmann transforms the state space to an equivalent continuous state space and uses the same condition. For a discrete state space case, there is an alternative characterization of Lehmann effectiveness and it is described in Appendix A.1.

[6] Note that $T(\omega; y)$ is nondecreasing in $y$ for sure by definition of $T$.

**Definition 2.4** (Monotone Likelihood Ratio Property)**.** An information structure $f$ is said to satisfy the monotone likelihood ratio property (MLRP) if and only if for all $x,\ x' \in X$ and $\omega,\ \omega' \in \Omega$ such that $x' > x$ and $\omega' > \omega$, the following inequality holds:

$$f(x|\omega) \cdot (x'|\omega') - f(x|\omega') \cdot f(x'|\omega) \geq 0.$$

The monotone likelihood ratio property means that a higher signal is more likely to be realized with a higher state, thus, quantiles of a signal $y$ of an information structure $g$, $\{G(y|\omega)\}_{\omega\in\Omega}$, convey the essential information about $y$. Then, by the definition, a set of signals, $\{T(\omega; y)\}_{\omega\in\Omega}$ of the information structure $f$, replicates $y$ of $g$ in the sense that they provide the same quantiles. The effectiveness condition means that a piece of information provided by only one signal in $g$ can be replicated by a nondecreasing sequence of signals in $f$. It implies that $f$ provides additional monotonic information about states after generating the same piece of information as in $g$, thus, $f$ can be more informative than $g$. This argument is largely based on the MLRP of information structures and it is well known that Lehmann's condition refines Blackwell's condition with this assumption. However, if the MLRP is relaxed especially on the supposedly higher information structure, Lehmann's condition no longer implies Blackwell's condition. The following proposition elaborates how Blackwell's condition and Lehmann's condition are related with and without the MLRP:

**Proposition 2.2.** *When $f$ and $g$ satisfy the MLRP, $f \geq_B g$ implies $f \geq_L g$, but $f \geq_L g$ does not imply $f \geq_B g$. When the MLRP of $f$ is not assumed, $f \geq_B g$ does not imply $f \geq_L g$.*

*Proof.* Jewitt [2007] showed that the first argument holds when $f$ and $g$ satisfy the MLRP. Consider information structures $f,\ g,\ h$ in Figure 2. Recall that $g$ and $h$ satisfy the MLRP, but $f$ does not. In Appendix A.2, I show that $g$ is more Lehmann effective than $h$ but not more Blackwell sufficient. In the appendix, it is also shown that $f$ is more Blackwell sufficient than $g$ but not more Lehmann effective. □

The next proposition shows that Lehmann effectiveness is weaker than monotone quasi-garbling by constructing generic state contingent noise.

**Proposition 2.3.** *$f \geq_L g$ implies $f \geq_{MQG} g$, but $f \geq_{MQG} g$ does not imply $f \geq_L g$.*

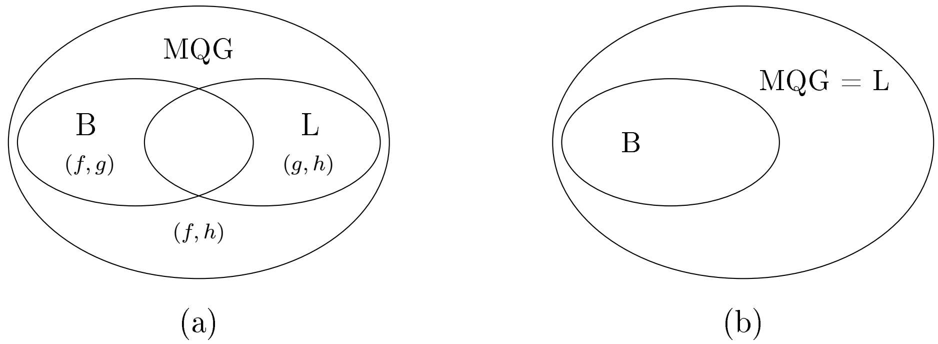


Figure 3: The relationship among Blackwell, Lehmann and Monotone Quasi-garbling - (a) When the MLRP is not assumed on the supposedly higher information structure, monotone quasi-garbling is weaker than both Blackwell and Lehmann, and Blackwell and Lehmann do not imply each other. Pairs of information structures $f$, $g$ and $h$ in Figure 2 are examples such that some conditions hold but the other conditions do not hold. (b) When the MLRP is assumed for both information structures, Blackwell is weaker than Lehmann and monotone quasi-garbling is equivalent to Lehmann.

*Proof.* Define $\gamma : X \times \Omega \to \Delta(Y)$ as follows:

$$\Gamma(y|x,\omega) = \begin{cases} 1 & \text{if } x \leq T(\omega; y), \\ 0 & \text{if } x > T(\omega; y). \end{cases} \tag{2.1}$$

Since $T(\omega; y)$ is nondecreasing in $y$ by the definition of $T$, for $y' > y$, $\Gamma(y|x,\omega) = 1$ implies that $\Gamma(y'|x,\omega) = 1$, thus, $\Gamma$ is a well defined as a cdf of a probability measure. Then, note that

$$G(y|\omega) = F(T(\omega; y)|\omega) = \int_{x \in X} \Gamma(y|x,\omega) dF(x|\omega), \tag{2.2}$$

thus, the constructed $\Gamma$ satisfies the first condition of monotone quasi-garbling.

From the definition of Lehmann effectiveness, $T(\omega; y)$ is nondecreasing in $\omega$. Then, for all $\omega' > \omega$ and $y \in Y$, we can see that

$$\Gamma(y|x,\omega') \geq \Gamma(y|x,\omega), \tag{2.3}$$

which implies that $\gamma(\cdot|x,\omega) \geq_{FOSD} \gamma(\cdot|x,\omega')$, therefore, $\gamma$ satisfies the second condition of monotone quasi-garbling.

Consider $f$ and $h$ in Figure 2. In Appendix A.2, I show that $f \geq_{MQG} h$ but

$f \ngeq_L h$. □

The next proposition establishes the equivalence between Lehmann effectiveness and monotone quasi-garbling under the assumption on the MLRP of the supposedly higher information structure. The proof is relegated to the appendix.

**Proposition 2.4.** *When $f$ satisfies the MLRP, $f \geq_L g$ if and only if $f \geq_{MQG} g$.*

Therefore, monotone quasi-garbling is generally weaker than Lehmann's condition, but equivalent to Lehmann's condition when the MLRP of $f$ is assumed. Recall that one of the purposes of this paper is to establish a criterion for monotone decision problems without imposing the MLRP on the supposedly higher information structure. In the main result (Theorem 1) of this paper, without imposing the MLRP on the supposedly higher information structure, it is shown that monotone quasi-garbling guarantees higher expected payoffs for monotone decision problems. Since monotone quasi-garbling is weaker than Lehmann's condition, the main result extends Lehmann's result.

The relationship among Blackwell, Lehmann, and monotone quasi-garbling criteria is summarized in Figure 3. Figure 3-(a) illustrates the relationship among Blackwell, Lehmann, and monotone quasi-garbling when the MLRP is not assumed especially on the supposedly higher information structure. On the other hand, if the MLRP is assumed for both information structures, the conditions are related as in Figure 3-(b).

I finalize the section by showing that monotone quasi-garbling is applicable in the example described in Figure 1.

**Proposition 2.5.** *If $f \geq_B g$ and $g \geq_L h$, then $f \geq_{MQG} h$.*[7]

*Proof.* From $f \geq_B g$, there exists $\gamma : \Omega \to \Delta(Y)$ that satisfies

$$g(y|\omega) = \int_X \gamma(y|x) f(x|\omega)$$

and let $\Gamma(\cdot|\omega)$ be the cdf of $\gamma(\cdot|\omega)$.

[7] This proposition is needed because transitivity of monotone quasi-garbling does not necessarily hold.

From $g \geq_L h$, $T(\omega; z) \equiv G^{-1}(H(z|\omega)|\omega)$ is nondecreasing in $\omega$. Define $\hat{\gamma} : X \times \Omega \to \Delta(Z)$ to be the pdf of $\hat{\Gamma}(z|x, \omega) \equiv \Gamma(T(\omega; z)|x)$. Then, $\{\hat{\gamma}(\cdot|x, \omega)\}_{\omega \in \Omega}$ is reversely FOSD-ordered due to $\hat{\Gamma}(z|x, \omega') = \Gamma(T(\omega'; z)|x) \geq \Gamma(T(\omega; z)|x) = \hat{\Gamma}(z|x, \omega)$ for $\omega' > \omega$.

Also note that

$$\begin{aligned} H(z|\omega) &= G(T(\omega; z)|\omega) \\ &= \int_X \Gamma(T(\omega; z)|x) f(x|\omega) dx = \int_X \hat{\Gamma}(z|x, \omega) f(x|\omega) dx. \end{aligned}$$

By differentiating the above equation by $z$, we can see that $h$ is obtained from $f$ by adding reversely monotone noise $\hat{\gamma}$. $\square$

# 3 Information in Monotone Decision Problems

In this section, I argue that monotone quasi-garbling can be a useful criterion for comparing information structures in monotone decision problems. I provide a general definition of monotone decision problems that can be applied even in multidimensional decision problems in Subsection 3.1. Then, I show that monotone quasi-garbling guarantees a higher expected payoff for any monotone decision problem in Subsection 3.2. It is also shown that monotone quasi-garbling is a necessary condition for informativeness for some cases in Subsection 3.3.

## 3.1 Monotone Decision Problems

The decision maker, hereafter DM, has a prior belief $\lambda \in \Delta(\Omega)$ and the cdf of $\lambda$ is given by $\Lambda$. When a prior belief is $\lambda$ and an information structure is $f$, the marginal signal distribution, $f_\lambda \in \Delta(X)$, would be

$$F_\lambda(x) = \int_{\omega \in \Omega} F(x|\omega) d\Lambda, \quad \text{and} \quad f_\lambda(x) = \frac{dF_\lambda(x)}{dx} = \int_{\omega \in \Omega} f(x|\omega) d\Lambda.$$

With a realized signal $x$, the posterior belief, $\lambda_f^x \in \Delta(\Omega)$, would be

$$\Lambda_f^x(\omega) = \frac{\int_{s \leq \omega} f(x|s) d\Lambda}{f_\lambda(x)}, \quad \text{and} \quad \lambda_f^x(\omega) = \frac{d\Lambda_f^x(\omega)}{d\omega} = \frac{f(x|\omega)\lambda(\omega)}{f_\lambda(x)}.$$

Similarly, define the above components for another information structure $g$ with the set of signals $Y \subset \mathbb{R}$. Let the set of alternatives or feasible decisions be $A \subseteq \mathbb{R}^n$ and DM's payoff function be given by $u : A \times \Omega \to \mathbb{R}$. Then, the decision making process is as follows:

0. DM has utility $u : A \times \Omega \to \mathbb{R}$ and a prior belief $\lambda$ on $\Omega$.

1. DM chooses an information structure between $f$ and $g$.

2. From the chosen info structure (say $f$), DM receives a signal $x \in X$.

3. DM updates a belief on the state to $\lambda_f^x(\omega)$ and chooses $a \in A$.

4. The state $\omega$ is revealed and the payoff $u(a, \omega)$ is realized.

In the third stage with a signal $x$, i.e., with posterior belief $\lambda_f^x \in \Delta(\Omega)$, DM chooses $a \in A$ to solve the following problem:

$$\max_{a \in A} \int_{\omega \in \Omega} u\left(a, \omega\right) d\Lambda_f^x, \tag{3.1}$$

and let the solution of the above problem be $a^*(x) \in A$. Then, the value of the information structure $f$ at the first stage is

$$\begin{aligned} V(f; u, \lambda) &= \mathbb{E}_X\left[\max_{a \in A} \int_{\omega \in \Omega} u\left(a, \omega\right) d\Lambda_f^x\right] \\ &= \int_{x \in X} \int_{\omega \in \Omega} u\left(a^*(x), \omega\right) d\Lambda_f^x dF_\lambda. \end{aligned} \tag{3.2}$$

I introduce two conditions which make a decision model monotone. In most decision models with monotonicity, it is assumed that the set of feasible decisions is single dimensional with a generic order and information structures satisfy the MLRP. To extend monotonicity to multidimensional decision sets $A \subseteq \mathbb{R}^n$, I impose a *well-ordered* condition, which corresponds to a generic order of single dimensional space, and a *monotonicity* condition, which corresponds to the MLRP of information structures.[8] I begin by defining the well-ordered condition.

[8]These correspondences will be elucidated in Subsection 4.1.

**Definition 3.1** (Well-Ordered Condition)**.** The set of feasible decisions $A$ satisfies the *well-ordered condition* with respect to $u$ if there exists a partial order $\geq_A$ such that for any decreasing (in terms of $\geq_A$) act $h : \Omega \to A$,[9] there exists $\hat{a} \in A$ such that

$$u(\hat{a}, \omega) \geq u(h(\omega), \omega) \quad \forall \omega \in \Omega. \tag{3.3}$$

A rank of alternatives, $\geq_A$, means that a highly ranked alternative is better for a higher state and worse for a lower state and vice versa for a lowly ranked alternative. According to this interpretation, a decreasing act should not be an optimal way of establishing an act because a highly ranked alternative (which is worse for a lower state) corresponds to a lower state and a lowly ranked alternative (which is worse for a higher state) corresponds to a higher state. Then, for any decreasing act, there should exist a feasible alternative $\hat{a} \in A$ that dominates the decreasing act and the inequality (3.3) ensures the existence of such alternative.[10]

The next step is to relate this rank of alternatives to the information structure. The Monotonicity condition guarantees that the optimal decision for the higher signal is ranked higher (in terms of the partial order $\geq_A$) than that for the lower signal.

**Definition 3.2** (Monotonicity Condition)**.** An information structure $g$ satisfies the *monotonicity condition* with respect to a tuple of a state contingent utility $u : A \times \Omega \to \mathbb{R}$, a prior belief $\lambda \in \Delta(\Omega)$, and a partial order $\geq_A$ on $A$, if the optimal policy $a^*(s)$ of the maximization problem in stage 3 satisfies

$$a^*(s') \geq_A a^*(s) \qquad \forall s' > s.$$

By the well-ordered and monotonicity conditions, monotone decision problems—a higher alternative would be chosen for a higher signal—can be formally defined. Informativeness is defined by comparing the value of information structures in the first stage for all monotone decision problems and is formally stated as follows:

[9] For $\omega' > \omega$ and $h : \Omega \to A$, $h$ is decreasing in terms of $\geq_A$ if $h(\omega) \geq_A h(\omega')$.

[10] This condition can be weakened by using a Bayesian approach rather than a dominance relation. To be specific, assume that for any posterior belief $v \in \Delta(\Omega)$ we can find an alternative $a(v) \in A$ such that yield higher expected utility than $h$:

$$\int_{\omega \in \Omega} u(a(v), \omega) v(\omega) d\omega \geq \int_{\omega \in \Omega} u(h(\omega), \omega) v(\omega) d\omega.$$

**Definition 3.3** (Informativeness for monotone decision problems)**.** An information structure $f$ is *more informative* than another information structure $g$ for monotone decision problems if and only if for all priors $\lambda \in \Delta(\Omega)$ and all payoff functions $u : \Omega \times A \to \mathbb{R}$ such that (i) $A$ satisfies the well-ordered condition with respect to $u$ with $\geq_A$; (ii) $g$ satisfies the monotonicity condition with respect to $(u, \lambda, \geq_A)$, the following inequality holds:

$$V(f; u, \lambda) \geq V(g; u, \lambda).$$

## 3.2 Monotone Quasi-garbling implies Informativeness

As discussed in Subsection 1.2, previous studies showed that Lehmann's effectiveness condition is a necessary and sufficient condition for informativeness when both information structures satisfy the MLRP and the utility function satisfies a monotone condition such as the single crossing property. Recall that monotone quasi-garbling is equivalent to Lehmann's condition under the MLRP of $f$, the supposedly higher information structure, so that monotone quasi-garbling is also a necessary and sufficient condition for informativeness in this class of problems.

One of the main questions of this paper is how to compare information structures without assuming the MLRP of the supposedly higher information structure. In the following theorem, it is shown that, even if the MLRP of $f$ is relaxed, monotone quasi-garbling is still a sufficient condition for informativeness for monotone decision problems.

**Theorem 1.** *If $g$ is a monotone quasi-garbling of $f$, $f$ is more informative than $g$ for monotone decision problems. In other words, when $\lambda \in \Delta(\Omega)$ is a prior belief, $A$ satisfies the well-ordered condition with respect to $u$ with $\geq_A$ and $g$ satisfies the monotonicity condition with respect to $(u, \lambda, \geq_A)$, and $g$ is a monotone quasi-garbling of $f$,*

$$V(f; u, \lambda) \geq V(g; u, \lambda).$$

Recall that monotone quasi-garbling is weaker than Lehmann's effectiveness condition when the MLRP of $f$ is not assumed. Therefore, the above theorem implies that Lehmann's effectiveness condition may no longer serve as a necessary condition for informativeness for monotone decision problems without the MLRP of $f$.

### 3.2.1 Sketch of Proof

Here I provide a sketch of the proof for Theorem 1 and relegate the detailed proof to Appendix B.2. One of the key steps for the proof is to transform a class of reversely FOSD-ordered probability measures on $Y$ ($\{\gamma(\cdot|\omega)\}_{\omega\in\Omega}$) (as in the second condition of the monotone quasi-garbling condition) into a probability measure on the set of nonincreasing functions from $\Omega$ to $Y$, denoted by $\mathcal{D} \subset Y^\Omega$. Recall that a decreasing act is used to define the well-ordered condition, thus, the transformation would link monotone decision problems and the monotone quasi-garbling condition.

First, consider $Y^\Omega$ as a Cartesian product $\prod_{\omega\in\Omega} Y$, and also consider a corresponding $\sigma$-algebra. Based on $\{\gamma(\cdot|\omega)\}_{\omega\in\Omega}$, we can define a probability measure on ($\hat{\gamma} \in \Delta(Y^\Omega)$) as follows: for any $h \in Y^\Omega$

$$\begin{aligned}\hat{\Gamma}(h) &= \Pr\left[k(\cdot) \in Y^\Omega \,:\, k(\omega) \leq h(\omega) \text{ for all } \omega \in \Omega\right] \\ &\equiv \inf_{\omega\in\Omega} \left\{\Gamma(h(\omega)|\omega)\right\},\end{aligned} \tag{3.4}$$

where $\hat{\Gamma}$ is the cdf of $\hat{\gamma}$. Note that this probability measure is well defined.[11] Then, the following lemma allows us to restrict attention to the set of nonincreasing functions $\mathcal{D} \subset Y^\Omega$:

**Lemma 3.1.** *Assume that $\{\gamma(\cdot|\omega)\}_{\omega\in\Omega}$ is reversely FOSD-ordered, i.e., $\gamma(\cdot|\omega) \geq_{FOSD} \gamma(\cdot|\omega')$ for all $\omega' > \omega$. If $h \notin \mathcal{D}$, i.e., $h(\omega) < h(\omega')$ for some $\omega' > \omega$, then, the probability measure $\hat{\gamma}$ assigns zero on $h$ ($\hat{\gamma}(h) = 0$).*

Now, $\hat{\gamma}$ can be considered as a probability measure on $\mathcal{D}$. The following lemma allows us to change a probability measure of an integration from $\gamma$ to $\hat{\gamma}$, i.e., from a series of reversely FOSD-ordered probability measures to a probability measure on decreasing functions.

[11] To verify that this probability measure is well defined, we need to check (i) the total measure is equal to one; (ii) probability is non-negative for any basis element in the sigma algebra. For the total measure, consider $\bar{h}(\omega) = \max Y$ for all $\omega$. Then, $\hat{\Gamma}(\bar{h})$ is the total probability and $\hat{\Gamma}(\bar{h}) = \inf\{\Gamma(\bar{h}(\omega)|\omega)\} = 1$. Note that any basis element $E$ in $\sigma$-algebra of $Y^\Omega$ can be represented by a difference between two functions $h$ and $h'$ with $h'(\omega) \geq h(\omega)$ for all $\omega \in \Omega$. Then, $\Gamma(h'(\omega)|\omega) \geq \Gamma(h(\omega)|\omega)$ for all $\omega \in \Omega$, thus, $\hat{\Gamma}(h') \geq \hat{\Gamma}(h)$. Since $\Pr(E) = \hat{\Gamma}(h') - \hat{\Gamma}(h)$, $\Pr(E)$ is non-negative.

**Lemma 3.2.** *For any* $u : A \times \Omega \to \mathbb{R}$, $a_g : Y \to A$, *reversely FOSD-ordered* $\{\gamma(\cdot|x,\omega)\}_{\omega\in\Omega}$ *and* $\omega \in \Omega$,

$$\int_{y\in Y} u(a_g(y),\omega)d\Gamma(y|x,\omega) = \int_{h\in\mathcal{D}} u(a_g(h(\omega)),\omega)d\hat{\Gamma}(h|x). \tag{3.5}$$

If $g$ is a monotone quasi-garbling of $f$, the value of the information structure $g$, $V(g;u,\lambda)$, would be represented as a weighted sum of the left hand side of (3.5) for all $x \in X$ where $a_g : Y \to A$ is the optimal decision for each signal of $g$. Then, by the monotonicity condition, $a_g$ is nondecreasing in $y$, thus, $a_g(h(\omega))$ is decreasing in $\omega$ for $h \in \mathcal{D}$. By the well-ordered condition, there exists an alternative $\phi(h) \in A$ that dominates $a_g(h(\cdot))$ and in the expectation given a signal $x \in X$, the optimal decision on $x$, $a_f(x)$, has to give a higher expected profit than $\phi(h)$. By integrating this for all $x \in X$, it can be shown that the value of the information structure $f$ is greater than that of $g$.

## 3.3 Necessary Condition of Informativeness for Special Cases

In this subsection, I explore the opposite direction of the implication of Theorem 1. Specifically, I show that monotone quasi-garbling is a necessary condition for informativeness for monotone decision problems in a restricted set of cases. Omitted proofs are provided in Appendix B.3.

### 3.3.1 Binary states

For the simplest case, suppose $\Omega = \{\omega_1, \omega_2\}$ with $\omega_1 < \omega_2$. In this case, $X$ can be ordered to satisfy the MLRP of $f$. To be specific, define an order $\geq$ on $X$ as follows:

$$x' \geq x \;\Leftrightarrow\; \frac{f(x'|\omega_2)}{f(x'|\omega_1)} \geq \frac{f(x|\omega_2)}{f(x|\omega_1)}.$$

Similarly, $Y$ can be ordered to satisfy the MLRP of $g$. Therefore, in this case, without assuming the MLRP of both information structures, it follows from $|\Omega| = 2$. Then, Lehmann's theorem is applicable and the effectiveness condition becomes a necessary condition of informativeness for monotone decision problems and so does monotone quasi-garbling by the following theorem.

**Theorem 2.** *If $|\Omega| = 2$ and $f$ is more informative than $g$ for monotone decision problems, then $f$ is more Lehmann effective and more Blackwell sufficient than $g$, thus, $g$ is a monotone quasi-garbling of $f$.*

*Proof.* By the above argument, when $|\Omega| = 2$, there exists orderings of $X$ and $Y$ that make $f$ and $g$ satisfy the MLRP. By Lehmann's theorem, when $f$ and $g$ satisfy the MLRP, effectiveness condition is a necessary condition for informativeness for decision problems with Karlin Rubin monotone functions. Since the class of decision problems is included in the monotone decision problems, Lehmann's condition becomes a necessary condition for informativeness for monotone decision problems. Then, we can apply Proposition 1 of Jewitt [2007] to show that when the state space is binary, Lehmann effectiveness is equivalent to Blackwell sufficiency.[12] By Proposition 2.1 (or 2.3), monotone quasi-garbling is a necessary condition for informativeness for monotone decision problems as well. □

Together with Theorem 1, when $|\Omega| = 2$, it is shown that monotone quasi-garbling is a necessary and sufficient condition for informativeness for monotone decision problems, moreover, so are Blackwell's and Lehmann's conditions. Therefore, all three criteria are equivalent for the binary state case.

**Corollary 3.1.** *When $|\Omega| = 2$, Blackwell's condition, Lehmann's condition, and monotone quasi-garbling are equivalent and they are necessary and sufficient conditions for informativeness for monotone decision problems.*

### 3.3.2 Binary signals for $f$

Suppose that $f$ has binary signals and $g$ satisfies a strict version of the MLRP. The following proposition shows that the strict MLRP of $f$ can be derived from informativeness for monotone decision problems.

**Proposition 3.2.** *If $|X| = 2$, $g$ satisfies the strict MLRP, and $f$ is more informative than $g$ for monotone decision problems, then $f$ satisfies the strict MLRP.*

[12] To be precise, Jewitt showed that Lehmann's condition is equivalent to Blackwell's 4th condition when $|\Omega| = 2$. The equivalence of Blackwell's 4th condition and Blackwell sufficiency condition in this paper is shown in Blackwell and Girshick [1954].

Based on this result, the effectiveness condition and the monotone quasi-garbling condition can be shown as a necessary condition for informativeness for monotone decision problems.

**Theorem 3.** *If $|X| = 2$, $g$ satisfies the strict MLRP, and $f$ is more informative than $g$ for monotone decision problems, then $f$ is more Lehmann effective than $g$, thus, $g$ is a monotone quasi-garbling of $f$.*

*Proof.* By Proposition 3.2, $f$ satisfies the strict MLRP. Then, when $f$ and $g$ satisfy the MLRP, since decision problems with Karlin Rubin monotone functions are included in monotone decision problems, $f$ is more Lehmann effective than $g$ by Lehmann's theorem. □

In this case, Lehmann's condition and monotone quasi-garbling become necessary and sufficient conditions for informativeness for monotone decision problems.

**Corollary 3.3.** *When $|X| = 2$ and $g$ satisfies the strict MLRP, Lehmann's condition and monotone quasi-garbling are equivalent and they are necessary and sufficient conditions for informativeness for monotone decision problems.*

### 3.3.3 Binary signals for $g$ and $|\Omega| = 3$

Suppose that $g$ has binary signals. In this case, the MLRP of $f$ is not acquired as a byproduct of assumptions or informativeness for monotone decision problems. Therefore, Lehmann's theorem cannot be applied as in the previous cases. The next theorem shows that monotone quasi-garbling is a necessary condition for informativeness for monotone decision problems when $|\Omega| = 3$ without using Lehmann's theorem.

**Theorem 4.** *If $|Y| = 2$, $|\Omega| = 3$, $g$ satisfies the MLRP, $f$ and $g$ have full support, and $f$ is more informative than $g$ for monotone decision problems, then $g$ is a monotone quasi-garbling of $f$.*

**Corollary 3.4.** *When $|Y| = 2$ and $|\Omega| = 3$, $g$ satisfies the MLRP, and $f$ and $g$ have full support, monotone quasi-garbling is a necessary and sufficient condition for informativeness for monotone decision problems.*

# 4 Applications

In this section, I present economic applications that illustrate monotone decision problems and monotone quasi-garbling as a useful way of comparing information structures. In Subsection 4.1, a single dimensional decision problem with the single crossing property will be rewritten in the grammar of Subsection 3.1. In the following subsection, I introduce an optimal insurance problem similar to the investment problem studied by (Cabrales et al. [2013]), and apply the main result of the paper to this setting. In Subsection 4.3, I consider a non-linear monopoly pricing problem.

## 4.1 Single dimensional decision with single crossing utility

Assume that DM's set of alternatives is single dimensional, i.e., $A_1 \subseteq \mathbb{R}$. Note that $A_1$ has a generic order. For preference, suppose that $u$ satisfies the single crossing property, that is, for any $a' > a$ and $\omega' > \omega$, $u(a', \omega) \geq u(a, \omega)$ implies $u(a', \omega') \geq u(a, \omega')$ and $u(a', \omega) > u(a, \omega)$ implies $u(a', \omega') > u(a, \omega')$.

The following lemma by Chi [2015] shows that $A_1$ satisfies the well-ordered condition w.r.t. $u$ with the generic order:

**Lemma 4.1.** *When $u$ satisfies the single crossing property, for any decreasing decision rule $h : \Omega \to A_1$, there exists $a \in A_1$ such that $u(a, \omega) \geq u(h(\omega), \omega)$ for all $\omega \in \Omega$.*

*Proof.* See Lemma 3.2 of Chi [2015]. □

Now assume that an information structure $f$ satisfies the MLRP. Then, by Athey [2002], when $u$ has the single crossing property and $f$ satisfies the MLRP, a higher alternative would be chosen for a higher signal. Therefore, DM's optimal choice would be nondecreasing in $x$ and the monotonicity condition would be satisfied.[13][14] Then, Theorem 1 can be applied as follows:

[13] In the monotone comparative statics result, the optimal decision is nondecreasing in $x$ in the sense of 'strong set order.' Hence, there is a chance that the realized decision is not actually nondecreasing. However, we can resolve this problem by assuming that DM chooses the greatest decision among the set of decisions that maximize the expected utility.

[14] The same argument is valid for a class of utility function satisfying the interval dominance order introduced by Quah and Strulovici [2009]. They present monotone comparative statics results for that class of utility functions. Chi [2015] shows Lemma 4.1 also holds for a class of utility functions that satisfies the interval dominance property as well (Lemma B.3).

**Proposition 4.1** (Single dimensional decision with single crossing utility)**.** *Assume that $A \subset \mathbb{R}$, $u : A \times \Omega \to \mathbb{R}$ satisfies the single crossing property and the decision maker solves*

$$\max\left(\arg\max_{a\in A}\int_{\omega\in\Omega} u(a,\omega)d\Upsilon\right),$$[15]

*in the third stage with a posterior belief $\Upsilon$. Then, if $g$ satisfies the MLRP and $g$ is a monotone quasi-garbling of $f$, then the decision maker gets greater expected utility in the first stage with $f$ than that with $g$ for all priors.*

## 4.2 An optimal insurance problem

An agent (DM) privately acquires information about a future state $\omega \in \Omega = \{\omega_1, \cdots, \omega_n\}$, which will be his realized income. Let $N = \{1, \cdots, n\}$ and $\omega_n > \cdots > \omega_1 > 0$. After receiving a signal, the agent can buy or sell Arrow securities. Assume that the Arrow security market is large enough so that the agent's demand does not affect prices. Let the price vector be $p = (p_1, \cdots, p_n)$ and assume that $\sum_{i=1}^n p_i = 1$ and $p_i > 0$.

In this example, an alternative would be the amount of securities bought for a state and say an alternative for $\omega_i$ as $q_i$. By allowing a short sale, the set of alternatives would be the set of real numbers. Assume that the agent is not allowed to execute a short sale more than $\omega_i$, i.e., $q_i + \omega_i \geq 0$ for all $i \in N$. Along with these constraints, the agent's budget constraint determines the set of feasible decisions. By assuming the agent's budget is zero, i.e., he must finance through short sales, we get the set of feasible decisions as follows:

$$A_2 = \{q = (q_1, \cdots, q_n) \in \mathbb{R}^n \,|\, p \cdot q \leq 0 \;\text{ and }\; \omega_i + q_i \geq 0 \;\; \forall i \in N\}. \tag{4.1}$$

Also note that the budget constraint will bind at the optimum and define the set of budget binding decisions as follows:

$$A_2' = \{q = (q_1, \cdots, q_n) \in \mathbb{R}^n \,|\, p \cdot q = 0 \;\text{ and }\; \omega_i + q_i \geq 0 \;\; \forall i \in N\}. \tag{4.2}$$

Let the agent's utility function be $u(q, \omega_i) = v(\omega_i + q_i)$ where $v$ is a CRRA utility function, i.e., $v(z) = z^{1-\rho}/(1-\rho)$ for some $\rho > 0$. In the third stage with the signal

[15]As mentioned in footnote 13, I require DM to choose the maximum alternative among the optimal alternatives in order to avoid the case that the decision is increasing in the strong set order but not actually increasing.

$x$, DM's problem is

$$\max_{q \in A_2} \sum_{i \in N} \lambda_f^x(\omega_i) v(\omega_i + q_i). \tag{4.3}$$

When an information structure is $\{f(\cdot|\omega_i)\}_{i \in N}$ and the agent receives $x$ as a signal, let the solution of (4.3) be $\{q_i(x)\}_{i \in N}$. Note that the solution can be obtained as follows:

$$\omega_i + q_i(x) = \left(\frac{\lambda_f^x(\omega_i)}{p_i}\right)^{1/\rho} \cdot \frac{\sum_{s=1}^n \omega_s \cdot p_s}{\sum_{s=1}^n \lambda_f^x(\omega_s)^{1/\rho} \cdot p_s{}^{1-1/\rho}}, \tag{4.4}$$

for all $i \in N$.

Define a partial order $\geq_{A_2'}$ on $A_2'$ as follows: $q' \geq_{A_2'} q$ if and only if

$$MRS'_{ij} \geq MRS_{ij} \;\Leftrightarrow\; MRS'_{ji} \leq MRS_{ji} \;\Leftrightarrow\; \frac{\omega_j + q'_j}{\omega_i + q'_i} \geq \frac{\omega_j + q_j}{\omega_i + q_i}, \tag{4.5}$$

for all $j > i$. Intuitively, if a lower state is likely to happen, DM would choose a lower bundle by $\geq_{A_2'}$ so that MRS for a lower state is low and that for a higher state is high.

Consider a decreasing act $\{q^j\}_{j \in N} \in (A_2')^n$ where $q^1 \geq_{A_2'} \cdots \geq_{A_2'} q^n$. Define $\hat{q}_i \equiv q_i^i$ and $\hat{q} = (\hat{q}_1, \cdots, \hat{q}_n)$. To check whether $A_2'$ satisfies the well-ordered condition w.r.t. $u$, it is enough to show that $\hat{q} \in A_2$.[16] Note that $\omega_i + \hat{q}_i = \omega_i + q_i^i \geq 0$ from $q^i \in A_2'$. Therefore, it suffices to show that $p \cdot \hat{q} \leq 0$.

For all $i, j \in N$, define $r_i^j \equiv p_i(\omega_i + q_i^j)/(\sum_{s=1}^n p_s \omega_s)$ and $r^j \equiv \{r_i^j\}_{i \in N}$. Note that $\sum_{i \in N} r_i^j = 1$ from $q^j \in A_2'$, thus, $r^j$ can be considered as an element of $\Delta(\Omega)$. Moreover, the condition (4.5) makes $\{r^j\}_{j \in N}$ reversely MLRP ordered, i.e., $r^1 \geq_{MLRP} \cdots \geq_{MLRP} r^n$. Therefore, $r^1 \geq_{FOSD} \cdots \geq_{FOSD} r^n$. Then, the following inequalities

[16] If $\hat{q} \in A_2$, we can find $\tilde{q} \in A_2'$ such that $\tilde{q}_i \geq \hat{q}_i$ for all $i \in N$, thus, $u(\tilde{q}, \omega_i) \geq u(\hat{q}, \omega_i) = u(q^i, \omega_i)$ for all $i \in N$.

hold:

$$\begin{aligned}
1 &= r_1^n + r_2^n + \cdots + r_{n-2}^n + r_{n-1}^n + r_n^n \\
&\geq (r_1^{n-1} + r_2^{n-1} + \cdots + r_{n-2}^{n-1} + r_{n-1}^{n-1}) + r_n^n && (\text{by } \ r^{n-1} \geq_{FOSD} r^n) \\
&\geq (r_1^{n-2} + r_2^{n-2} + \cdots + r_{n-2}^{n-2}) + r_{n-1}^{n-1} + r_n^n && (\text{by } \ r^{n-2} \geq_{FOSD} r^{n-1}) \\
&\ \ \vdots \\
&\geq r_1^1 + r_2^2 + \cdots r_{n-2}^{n-2} + r_{n-1}^{n-1} + r_n^n && (\text{by } \ r^1 \geq_{FOSD} r^2) \\
&= \sum_{i \in N} \frac{p_i(\omega_i + \hat{q}_i)}{\sum_{s=1}^n p_s \omega_s}
\end{aligned}$$

From the last inequality, $0 \geq \sum_{i \in N} p_i \hat{q}_i = p \cdot \hat{q}$. Therefore, $A_2'$ satisfies the well-ordered condition w.r.t. $u$.

Assume that an information structure $\{f(\cdot|\omega)\}_{\omega \in \Omega}$ satisfies the MLRP. By Milgrom [1981], the posterior belief also shows the MLRP, that is, for all $x' > x$ and $j > i$,

$$\lambda_f^{x'}(\omega_j)\lambda_f^x(\omega_i) - \lambda_f^x(\omega_j)\lambda_f^{x'}(\omega_i) \geq 0 \quad \Leftrightarrow \quad \frac{\lambda_f^{x'}(\omega_j)}{\lambda_f^x(\omega_j)} \geq \frac{\lambda_f^{x'}(\omega_i)}{\lambda_f^x(\omega_i)}. \tag{4.6}$$

We see that this condition represents the intuition that with a lower signal, a lower state is more likely to happen and with a higher signal, a higher state is more likely to happen. Note that from (4.4),

$$\frac{\omega_j + q_j(x')}{\omega_j + q_j(x)} \geq \frac{\omega_i + q_i(x')}{\omega_i + q_i(x)} \quad \Leftrightarrow \quad \frac{\lambda_f^{x'}(\omega_j)^{1/\rho}}{\lambda_f^x(\omega_j)^{1/\rho}} \geq \frac{\lambda_f^{x'}(\omega_i)^{1/\rho}}{\lambda_f^x(\omega_i)^{1/\rho}}, \tag{4.7}$$

and it holds from (4.6). Therefore, we can see that

$$\{q_i(x')\}_{i \in N} \geq_{A_2'} \{q_i(x)\}_{i \in N}$$

for all $x' > x$ and the monotonicity condition holds. Then, Theorem 1 can be applied as follows:

**Proposition 4.2** (Optimal insurance problem with the MLRP)**.** *In the optimal insurance problem, when $g$ satisfies the MLRP and $g$ is a monotone quasi-garbling of $f$, then the decision maker obtains greater expected utility in the first stage from $f$*

*than from g for all priors.*[17]

## 4.3 Nonlinear Pricing under Monopoly

As a final application of monotone decision problems, I investigate information ranking in the context of monopoly pricing with second-degree price discrimination as in Maskin and Riley [1984]. As noted in Subsection 1.2, comparison of information structures for monopoly pricing has been studied by Athey and Levin [2017] and Ottaviani and Prat [2001]. Athey and Levin applied their result in a setting where the seller receives a signal about the buyer's type and determines the simple monopoly price and there is no second-degree price discrimination. In the study by Ottaviani and Prat, both the seller and the buyer receive a signal about the buyer's type and the seller offers a nonlinear price. In this subsection, I study the mixture of these two models: the buyer is fully aware of her type, the seller receives a signal about the buyer's type and then a menu of nonlinear prices is offered.

Thus, there is an information asymmetry between the seller and the buyer about the buyer's type, and the menu of prices and quantities maximizes expected monopoly profit subject to the incentive compatibility constraints (ICC) and the individual rationality constraints (IRC). Although this is usually modeled as a two-player game, it can also be cast as a multidimensional decision problem for a seller constrained by ICC and IRC by regarding the buyer's type as a state. Moreover, with fairly mild assumptions, this application satisfies the well-ordered and monotonicity conditions so that the main result is applicable.

### 4.3.1 Model and the Basic Results

In this subsection, I review the model and the basic results of Maskin and Riley [1984] and rewrite the model in the grammar of Section 3.

A monopolistic seller (DM) sells a good to a buyer. The buyer has a type $\omega \in \Omega = [0, 1]$, which is known to the buyer, but not the seller. Let the utility of a type $\omega$ buyer from consuming $q$ units of good and paying $p$ units of money be $v(q, \omega) - p$. Let the cost function of the seller be $c : \mathbb{R}_+ \to \mathbb{R}_+$, then the seller's payoff would be

[17] When the decision maker's prior is known, the entropy criterion by Cabrales et al. [2013] gives a complete ordering. The result of this proposition is, therefore, useful when the decision maker's prior is unknown.

$u(q, p, \omega) = -c(q) + p$.[18] Note that without information asymmetry, for all $\omega \in \Omega$, the seller would offer the first best allocation $q^*(\omega)$ such that $c'(q^*(\omega)) = v_q(q^*(\omega), \omega)$. Assume some conditions on $c$ and $v$.

**Assumption.** The seller's cost function $c(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ and the buyer's utility function $v : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ satisfy the following properties:

1. $c$ is twice differentiable and convex in $q$ ($c'' \geq 0$),
2. $v$ is three times differentiable ($v_{qq\omega}$ exists),
3. $v$ is concave in $q$ and supermodular in $(q, \omega)$ ($v_{qq} \leq 0$, $v_{q\omega} \geq 0$),
4. $v_q$ is supermodular in $(q, \omega)$ ($v_{qq\omega} \geq 0$),
5. $c(0) = 0$ and $v(q, 0) = 0$ for all $q \geq 0$.

We also assume that the seller possesses all the bargaining power (i.e., moves first in the game) and posts a menu of non-linear pricing options from which the buyer must select if she wishes to transact. The seller has a prior belief $\Lambda \in \Delta(\Omega)$. After receiving a signal, the seller forms a posterior belief on the buyer's type and sets pricing options $\{q(\omega), p(\omega)\}_{\omega \in \Omega}$ based on the belief. The set of alternatives can be viewed as a subset of

$$(\mathbb{R}^2_+)^\Omega \equiv \left\{ \{q(\cdot), p(\cdot)\} \mid \{q(\cdot), p(\cdot)\} : \Omega \to \mathbb{R}^2_+ \right\}.$$

To implement a pricing option, it must satisfy the following constraints:

$$v(q(\omega), \omega) - p(\omega) \geq v(q(\omega'), \omega) - p(\omega'), \quad \forall \omega' \in \Omega, \tag{ICC}$$
$$v(q(\omega), \omega) - p(\omega) \geq 0 \quad \forall \omega' \in \Omega. \tag{IRC}$$

Therefore, the set of feasible decisions is

$$A_3 = \left\{ \{q(\cdot), p(\cdot)\} \in \mathbb{R}^\Omega_+ \,\middle|\, \{q(\cdot), p(\cdot)\} \text{ satisfies ICCs and IRCs} \right\}. \tag{4.8}$$

[18] Note that in this example, DM's utility itself is not dependent on $\omega$. However, different payoffs across states come from DM's price discrimination.

Then, when the posterior belief is $\Upsilon(\cdot)$ and $\Upsilon$ has full support,[19] DM's problem is

$$\max_{\{q(\cdot),p(\cdot)\}\in A_3} \int_{\omega\in\Omega} \left[-c(q(\omega)) + p(\omega)\right] d\Upsilon(\omega). \tag{4.9}$$

It is well known that an allocation $\{q(\cdot)\}_{\omega\in\Omega}$ is implementable iff $q(\cdot)$ is monotone, and all the local downward ICCs and IRC for the lowest type are binding at the solution of (4.9) (Maskin and Riley [1984], Guesnerie and Laffont [1984]). Therefore, the solution of (4.9) would be in the set of decisions $A_3' \subset A_3$ defined as follows:

$$A_3' = \left\{ \{q(\cdot),p(\cdot)\} \in \mathbb{R}_+^\Omega \;\middle|\; \begin{array}{l} q(\omega') \geq q(\omega) \text{ for all } \omega' > \omega, \\ p(\omega) = \int_0^\omega v_q(q(s),s)q'(s)ds, \end{array} \right\}.^{20} \tag{4.10}$$

By restricting (4.9) to $A_3'$, we can rewrite (4.9) as follows:

$$\max_{\text{nondecreasing } q(\cdot)} \int_{\omega\in\Omega} I(q(\omega),\omega)d\Upsilon(\omega), \tag{4.11}$$

where

$$I(q,\omega) = -c(q) + v(q,\omega) - v_\omega(q,\omega)\cdot\frac{1-\Upsilon(\omega)}{\upsilon(\omega)}. \tag{4.12}$$

Let $\bar{q}(\omega)$ be the second best allocation, which is the solution of $\frac{\partial I}{\partial q}(q,\omega) = 0$, thus,

$$c'(\bar{q}(\omega)) = v_q(\bar{q}(\omega),\omega) - v_{q\omega}(\bar{q}(\omega),\omega)\cdot\frac{1-\Upsilon(\omega)}{\upsilon(\omega)}. \tag{4.13}$$

From $c'' \geq 0$ and $v_{qq} \leq 0$, $v_q(q,\omega) - c'(q) \geq 0$ for all $q \leq q^*(\omega)$ and also note that $\bar{q}(\omega) \leq q^*(\omega)$.

If $\bar{q}(\omega)$ is nondecreasing in $\omega$, we can easily see that the optimal solution is $\bar{q}(\cdot)$. However, if there is a decreasing part in $\bar{q}(\cdot)$, we need to utilize the 'ironing' technique.[21] See Appendix B.4 for the full characterization of the solution to (4.11).

[19] If $\Upsilon$ does not have full support, the seller would ignore some ICCs and IRCs and suggest a menu of pricing which may not be in $A_3$.

[20] Note that the IRC for the lowest type binds so that $p(0) = u(q(0),0) = 0$.

[21] The ironing technique originated in Myerson [1981] and was extended to general objective functions by Toikka [2011].

### 4.3.2 Monotonicity and Information Ranking

Under the optimal mechanism, no matter what the seller's belief is, quantities for higher types are almost as high as the first best quantity. However, quantities for lower types serve as a tool for incentivizing not only the lower types of the buyer but also higher types through the pricing policy $p(\omega) = \int_0^\omega v_q(q(s), s)q'(s)ds$. From $v_{qq} \leq 0$, we can see that a lower quantity induces a higher price for a higher type buyer. Note that if the seller believes that the buyer is more likely to be a high type, then the seller would be less concerned about the possibility of distorting the quantities designed for low types in order to raise the prices for high types.

Using this intuition, consider a partial order $\geq_{A'_3}$ on $A'_3$ defined as follows: for any $\{q(\cdot), p(\cdot)\}$, $\{q'(\cdot), p'(\cdot)\} \in A'_3$, $\{q'(\cdot), p'(\cdot)\} \geq_{A'_3} \{q(\cdot), p(\cdot)\}$ if and only if

$$q(\omega) \geq q'(\omega) \tag{4.14}$$

for all $\omega \in \Omega$. The next lemma justifies this partial order by showing that it satisfies the well-ordered condition.

**Lemma 4.2.** *When $v_{qq} \leq 0$, $u(q, p, \omega) = -c(q) + p$, and $A'_3$ is defined by (4.10), $A'_3$ satisfies the well-ordered condition with respect to $u$ with a partial order $\geq_{A'_3}$ defined by (4.14).*

The next task is to check the monotonicity condition. When an information structure is $\{f(\cdot|\omega)\}_{\omega\in\Omega}$ and DM receives $x$ as a signal, let the solution of (4.9) be $\{q_f(\omega; x), p_f(\omega; x)\}_{\omega\in\Omega}$, and the second best solution be $\{\bar{q}_f(\omega; x), \bar{p}_f(\omega; x)\}_{\omega\in\Omega}$. The following result shows that the monotonicity condition holds with assumptions 1-5 and the MLRP of an information structure.

**Lemma 4.3.** *Suppose that assumptions 1-5 hold for $(v, c)$ and an information structure $f$ satisfies the MLRP and has full support for all $\omega$. Also assume that $u(q, p, \omega) = -c(q) + p$, $A'_3$ is defined by (4.10), $\geq_{A'_3}$ is defined by (4.14), and DM solves (4.11). Then, for any prior $\lambda$ with full support, $f$ satisfies the monotonicity condition with respect to $(u, \lambda, \geq_{A'_3})$.*

I conclude this subsection by applying the main result of the paper in this framework.

**Proposition 4.3.** *Suppose that assumptions 1-5 hold for $(v, c)$, $f$, $g$ have full support for all $\omega$, $g$ satisfies the MLRP, and $g$ is a monotone quasi-garbling of $f$. Then, the seller prefers $f$ to $g$ for any prior belief $\lambda$ with full support.*

*Proof.* By Lemma 4.3, $g$ satisfies the monotonicity condition with respect to $u$, any prior belief $\lambda$ with full support, and $\geq_{A_3'}$. Then, by Theorem 1, $V(f; u, \lambda) \geq V(g; u, \lambda)$. $\square$

# 5 Conclusion

In this paper, I introduced the concept of monotone quasi-garbling that can serve as a prior independent criterion for monotone decision problems. This new ordering allows the comparison of a pair of information structures that are not comparable by Blackwell's or Lehmann's conditions, especially when the supposedly higher information structure does not satisfy the MLRP. Moreover, this new ordering provides an intuitive interpretation on information comparison in monotone decision problems: if an information structure is generated by adding a reversely monotone noise to the original information structure, it yields lower expected payoffs for monotone decision problems than those of the original one.

General formulation of monotone decision problems using the well-ordered and monotonicity conditions enlarges the class of problems that can be analyzed. The result can be applied to economic situations that entail multidimensional or state contingent decisions such as optimal insurance or nonlinear monopoly pricing.

As for future work toward a characterization, although some special cases were covered, the necessary condition for informativeness for monotone decision problems in general is yet to be discovered.

# A Appendix: Discrete Setup

## A.1 Lehmann Effectiveness in Discrete Setup

In this appendix, I characterize Lehmann effectiveness in the discrete case by using some matrices. Assume that the state space is $\Omega = \{\omega_1, \cdots, \omega_n\}$ and signal spaces of $f$ and $g$ are $X = \{x_1, \cdots, x_K\}$ and $Y = \{y_1, \cdots, y_L\}$.

To properly define Lehmann effectiveness in the discrete type model, let us introduce some notions. Define $1 \leq \tilde{k}(y_l, \omega_i) \leq K$ as the unique index satisfying

$$F(x_{\tilde{k}(y_l,\omega_i)}|\omega_i) \leq G(y_l|\omega_i) < F(x_{\tilde{k}(y_l,\omega_i)+1}|\omega_i),$$

and also define $c(y_l, \omega_i)$ as the value satisfying

$$G(y_l|\omega_i) = F(x_{\tilde{k}(y_l,\omega_i)}|\omega_i) + c(y_l, \omega_i) \cdot f(x_{\tilde{k}(y_l,\omega_i)+1}|\omega_i),$$

and $0 \leq c(y_l, \omega_i) < 1$. Now we are ready to define Lehmann effectiveness in the discrete setup.

**Definition.** In the discrete setup, an information structure $f$ is more Lehmann effective than $g$ if, for all $\omega_i < \omega_j$ and $1 \leq l \leq L$,

1. $\tilde{k}(y_l, \omega_i) \leq \tilde{k}(y_l, \omega_j)$,
2. $c(y_l, \omega_i) \leq c(y_l, \omega_j)$ when $\tilde{k}(y_l, \omega_i) = \tilde{k}(y_l, \omega_j)$.

This definition is in line with continuous setup in that a pair $(\tilde{k}(\cdot, \omega_i), c(\cdot, \omega_i))$ plays a role as an inverse function of $F(\cdot|\omega_i)$. By introducing a notation, an easy way to check whether an information structure is more Lehman effective than another information structure or not is developed (Proposition A.1). I introduce a matrix that relates two information structures.

**Definition.** A matrix,

$$T_f^g(\omega_i) = \begin{pmatrix} a_{11}(\omega_i) & \cdots & a_{1L}(\omega_i) \\ \vdots & \ddots & \vdots \\ a_{K1}(\omega_i) & \cdots & a_{KL}(\omega_i) \end{pmatrix},$$

is said to be the **monotone transforming matrix** from $f$ to $g$ for $\omega_i$ if it satisfies following properties:

1. $g(y|\omega_i) = f(x|\omega_i) \cdot T_f^g(\omega_i)$,
2. $a_{kl}(\omega_i) > 0$ for $l \in [s_k(\omega_i), t_k(\omega_i)]$, and $a_{kl}(\omega_i) = 0$ otherwise,
3. $s_1(\omega_i) = 1$, $s_{k+1}(\omega_i) \in \{t_k(\omega_i), t_k(\omega_i) + 1\}$, $t_K(\omega_i) = L$,

4. $\sum_{l=s_k(\omega_i)}^{t_k(\omega_i)} a_{kl}(\omega_i) = 1$ for all $1 \leq k \leq K$.

**Example.** A pair of information structures $f$, $g$ (with $K = L = 4$) and the monotone transforming matrices that relate two information structures:

$$\begin{pmatrix} g(y_1|\omega_1) \\ g(y_2|\omega_1) \\ g(y_3|\omega_1) \\ g(y_4|\omega_1) \end{pmatrix}^{\intercal} = \begin{pmatrix} f(x_1|\omega_1) \\ f(x_2|\omega_1) \\ f(x_3|\omega_1) \\ f(x_4|\omega_1) \end{pmatrix}^{\intercal} \cdot \begin{pmatrix} .8 & .2 & 0 & 0 \\ 0 & .7 & .3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} g(y_1|\omega_2) \\ g(y_2|\omega_2) \\ g(y_3|\omega_2) \\ g(y_4|\omega_2) \end{pmatrix}^{\intercal} = \begin{pmatrix} f(x_1|\omega_2) \\ f(x_2|\omega_2) \\ f(x_3|\omega_2) \\ f(x_4|\omega_2) \end{pmatrix}^{\intercal} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ .1 & .7 & .2 & 0 \\ 0 & 0 & .3 & .7 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} g(y_1|\omega_3) \\ g(y_2|\omega_3) \\ g(y_3|\omega_3) \\ g(y_4|\omega_3) \end{pmatrix}^{\intercal} = \begin{pmatrix} f(x_1|\omega_3) \\ f(x_2|\omega_3) \\ f(x_3|\omega_3) \\ f(x_4|\omega_3) \end{pmatrix}^{\intercal} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ .3 & .5 & .2 & 0 \\ 0 & 0 & .7 & .3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that the monotone transforming matrices can be regarded as the set of probability measures that satisfy the first condition of the monotone quasi-garbling relation. Then, the following proposition shows that the Lehmann effectiveness requires this specific probability measures to satisfy the second condition of monotone quasi-garbling.

**Proposition A.1.** *In the discrete setup, an information structure $f$ is more Lehmann effective than $g$ if and only if*

$$\sum_{l \leq l'} a_{kl}(\omega_i) \leq \sum_{l \leq l'} a_{kl}(\omega_j),$$

*for all $1 \leq k \leq K$, $1 \leq l' \leq L$ and $\omega_i < \omega_j$, where $a_{kl}(\omega_i)$ is a component of monotone transforming matrix $T_f^g(\omega_i)$.*

*Proof.* From the definition of $T_f^g(\omega_i)$ and $g(y_{l'}|\omega_i)$, we can derive that

$$G(y_{l'}|\omega_i) = \sum_{k=1}^{K} f(x_k|\omega_i) \cdot \left( \sum_{l=1}^{l'} a_{kl}(\omega_i) \right),$$

and $\sum_{l=1}^{l'} a_{kl}(\omega_i)$ is equal to

1. 1, when $k \leq \tilde{k}(y_{l'}, \omega_i)$,

2. $c(y_{l'}, \omega_i)$, when $k = \tilde{k}(y_{l'}, \omega_i) + 1$,

3. 0, when $k > \tilde{k}(y_{l'}, \omega_i) + 1$.

Then, $f$ is more Lehmann effective than $g$, for $\omega_i < \omega_j$ and $1 \leq l' \leq L$ if and only if

1. for $k \leq \tilde{k}(y_l, \omega_i) \leq \tilde{k}(y_l, \omega_j)$,

$$1 = \sum_{l \leq l'} a_{kl}(\omega_i) \leq \sum_{l \leq l'} a_{kl}(\omega_j) = 1,$$

2. for $k = \tilde{k}(y_l, \omega_i) + 1 = \tilde{k}(y_l, \omega_j) + 1$,

$$c(y_l, \omega_i) = \sum_{l \leq l'} a_{kl}(\omega_i) \leq \sum_{l \leq l'} a_{kl}(\omega_j) = c(y_l, \omega_j),$$

3. for $k = \tilde{k}(y_l, \omega_i) + 1 < \tilde{k}(y_l, \omega_j) + 1$,

$$c(y_l, \omega_i) = \sum_{l \leq l'} a_{kl}(\omega_i) \leq \sum_{l \leq l'} a_{kl}(\omega_j) = 1,$$

4. for $\tilde{k}(y_l, \omega_i) + 1 < k$,

$$0 = \sum_{l \leq l'} a_{kl}(\omega_i) \leq \sum_{l \leq l'} a_{kl}(\omega_j).$$

□

Through this proposition, we can see the relationship between Lehmann effectiveness and monotone quasi-garbling again. Monotone quasi-garbling only requires the existence of a set of probability measures satisfying regeneration of an information structure (condition 1) and a reversely FOSD-order (condition 2). However, the Lehmann effectiveness requires a specific set of probability measures to satisfy a reversely FOSD-order, thus, the Lehmann effectiveness is stronger than monotone quasi-garbling. Now, we provide a specific example which shows that the Lehmann effectiveness is strictly stronger than monotone quasi-garbling.

## A.2 Details for the example in Figure 2

Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and information structures of Fred ($f$), George ($g$) and Hannah ($h$) are given as 2. In this subsection, I will show how their information structures are related by Blackwell, Lehmann, and monotone quasi-garbling criteria. Before establishing the relationship, consider a special case such that a prior $\gamma$ is uniformly distributed and utility function $u : A \times \Omega \to \mathbb{R}$ with $A = \{a_L, a_H\}$ is given as follows:

$$u(a|\omega) = \begin{array}{ccc} & a_L & a_H \\ \omega_1 & 1 & 0 \\ \omega_2 & 0 & 1 \\ \omega_3 & 0 & 1 \\ \omega_4 & 1 & 0 \end{array}. \tag{A.1}$$

Then, the optimized expected utility for each information structure would be

$$V(h; u, \lambda) = \frac{13}{24} \geq \frac{1}{2} = V(g; u, \lambda) = V(f; u, \lambda). \tag{A.2}$$

Note that $u$ is not a single crossing utility, moreover, any ordering cannot satisfy well-ordered condition.[22] Therefore, this case will only be used in checking Blackwell relation.

[22]If $a_H >_A a_L$, a decreasing function $h$ with $h(\omega_3) = a_H$ and $h(\omega_4) = a_L$ is not dominated by any alternative in $A$. If $a_L > a_H$, a decreasing function $h'$ with $h'(\omega_1) = a_L$ and $h'(\omega_2) = a_H$ is not dominated by any alternative in $A$.

### A.2.1 $f$ vs. $g$

1. ($f \geq_B g$) $f$ is more Blackwell sufficient than $g$ by

$$g = f \cdot \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

2. ($f \not\gtrsim_L g$) $f$ is *not* more Lehmann effective than $g$. To see this, check monotone transforming matrices:

$$T_f^g(\omega_1) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad T_f^g(\omega_2) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad T_f^g(\omega_3) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad T_f^g(\omega_4) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

   Note that $\{T_f^h(\omega)|_{x=LH}\}_{\omega \in \Omega}$ is not reversely FOSD-ordered so that $f$ is not more effective than $h$ in Lehmann's condition.

3. ($f \geq_{MQG} g$) $g$ is a monotone quasi-garbling of $f$ since $g$ is a garbling of $f$.

### A.2.2 $g$ vs. $h$

1. ($g \not\gtrsim_B h$) $g$ is *not* more Blackwell sufficient than $h$ due to (A.2).[23]

2. ($g \geq_L h$) $g$ is more Lehmann effective than $h$. To see this, check monotone transforming matrices:

$$T_g^h(\omega_1) = \begin{pmatrix} 5/6 & 1/6 \\ 0 & 1 \end{pmatrix}, \qquad T_g^h(\omega_2) = \begin{pmatrix} 1 & 0 \\ 1/3 & 2/3 \end{pmatrix},$$
$$T_g^h(\omega_3) = \begin{pmatrix} 1 & 0 \\ 1/3 & 2/3 \end{pmatrix}, \qquad T_g^h(\omega_4) = \begin{pmatrix} 1 & 0 \\ 1/3 & 2/3 \end{pmatrix}.$$

[23] If $g$ is more Blackwell sufficient than $h$, for all utility and prior, $V(h; u, \lambda)$ has to be greater than $V(g; u, \lambda)$.

Note that, for both $z = L, H$, $\{T_g^h(\omega)|_z\}_{\omega\in\Omega}$ is reversely FOSD-ordered so that $f$ is not more effective than $h$ in Lehmann's condition.

3. ($g \geq_{MQG} h$) $h$ is a monotone quasi-garbling of $g$ since $g$ is more effective than $h$ in Lehmann's condition.

### A.2.3 $f$ vs. $h$

1. $f$ is *not* more Blackwell sufficient than $h$ due to (A.2).

2. $f$ is *not* more Lehmann effective than $h$. To see this, check monotone transforming matrices:

$$T_f^h(\omega_1) = \begin{pmatrix} 1 & 0 \\ \frac{2}{3} & \frac{1}{3} \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad T_f^h(\omega_2) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \\ 0 & 1 \end{pmatrix}, \quad T_f^h(\omega_3) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}, \quad T_f^h(\omega_4) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ \frac{2}{3} & \frac{1}{3} \\ 0 & 1 \end{pmatrix}.$$

Note that $\{T_f^h(\omega)|_{x=LH}\}_{\omega\in\Omega}$ is not reversely FOSD-ordered so that $f$ is not more Lehmann effective than $h$.

3. $h$ is a monotone quasi-garbling of $f$ by following matrices:

$$\gamma_f^h(\omega_1) = \begin{pmatrix} 1 & 0 \\ \frac{2}{3} & \frac{1}{3} \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad \gamma_f^h(\omega_2) = \begin{pmatrix} 1 & 0 \\ \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \\ 0 & 1 \end{pmatrix}, \quad \gamma_f^h(\omega_3) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}, \quad \gamma_f^h(\omega_4) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

# B Appendix: Proofs

## B.1 Proofs for Section 2

*Proof for Proposition 2.4.* Assume that there exists $\omega' > \omega$ and $y \in Y$ such that $T(\omega'; y) < T(\omega; y)$. By rewriting (2.2), we can get

$$\begin{aligned} 1 &= \frac{\int_{x>T(\omega;y)} \Gamma(y|x,\omega) dF(x|\omega)}{\int_{x\leq T(\omega;y)}(1-\Gamma(y|x,\omega))dF(x|\omega)} \\ &= \frac{\int_{x>T(\omega';y)} \Gamma(y|x,\omega') dF(x|\omega')}{\int_{x\leq T(\omega';y)}(1-\Gamma(y|x,\omega'))dF(x|\omega')}. \end{aligned} \tag{B.1}$$

Note that

$$\frac{\int_{x>T(\omega;y)} \Gamma(y|x,\omega) dF(x|\omega)}{\int_{x\leq T(\omega;y)}(1-\Gamma(y|x,\omega))dF(x|\omega)} \leq \frac{\int_{x>T(\omega;y)} \Gamma(y|x,\omega) dF(x|\omega')}{\int_{x\leq T(\omega;y)}(1-\Gamma(y|x,\omega))dF(x|\omega')}, \tag{B.2}$$

from double integration of the MLRP conditions multiplied by $\Gamma(y|x,\omega)(1-\Gamma(y|x,\omega))$.

Then, from the condition 2 of monotone quasi-garbling, $\Gamma(y|x,\omega') \geq \Gamma(y|x,\omega)$ and $1-\Gamma(y|x,\omega) \geq 1-\Gamma(y|x,\omega')$ and from this, we can get

$$\frac{\int_{x>T(\omega;y)} \Gamma(y|x,\omega) dF(x|\omega')}{\int_{x\leq T(\omega;y)}(1-\Gamma(y|x,\omega))dF(x|\omega')} \leq \frac{\int_{x>T(\omega;y)} \Gamma(y|x,\omega') dF(x|\omega')}{\int_{x\leq T(\omega;y)}(1-\Gamma(y|x,\omega'))dF(x|\omega')}. \tag{B.3}$$

Lastly, from $T(\omega'; y) < T(\omega; y)$, we can get

$$\frac{\int_{x>T(\omega;y)} \Gamma(y|x,\omega') dF(x|\omega')}{\int_{x\leq T(\omega;y)}(1-\Gamma(y|x,\omega'))dF(x|\omega')} \leq \frac{\int_{x>T(\omega';y)} \Gamma(y|x,\omega') dF(x|\omega')}{\int_{x\leq T(\omega';y)}(1-\Gamma(y|x,\omega'))dF(x|\omega')}. \tag{B.4}$$

Note that the equality of (B.4) holds iff $\Gamma(y|x,\omega') = 1-\Gamma(y|x,\omega') = 0$ for all $T(\omega; y) \leq x \leq T(\omega'; y)$ and it is impossible to happen. Therefore, (B.4) holds for strict inequality.

By combining (B.2), (B.3), and (B.4), we get

$$\frac{\int_{x>T(\omega;y)} \Gamma(y|x,\omega) dF(x|\omega)}{\int_{x\leq T(\omega;y)}(1-\Gamma(y|x,\omega))dF(x|\omega)} < \frac{\int_{x>T(\omega';y)} \Gamma(y|x,\omega') dF(x|\omega')}{\int_{x\leq T(\omega';y)}(1-\Gamma(y|x,\omega'))dF(x|\omega')},$$

which is contradiction to (B.1). Therefore, with the MLRP of $f$, if $g$ is a monotone quasi-garbling of $f$, $f$ is more Lehmann effective than $g$. $\square$

## B.2 Proofs for Subsection 3.2

*Proof of Lemma 3.1.* Consider a function $l$ such that $l(s) = h(s)$ for all $s \neq \omega'$ and $l(\omega') = h(\omega) < h(\omega')$. Then, from reversely FOSD relation of $\gamma$, we can get

$$\Gamma(h(\omega)|\omega) = \Gamma(l(\omega)|\omega) \leq \Gamma(l(\omega')|\omega') \leq \Gamma(h(\omega')|\omega').$$

Then, we can see that $\omega'$ is ruled out as a minimizer of (3.4) for both $h$ and $l$.

Since $h$ and $l$ are equal for all states except for $\omega'$, $\hat{\Gamma}(h) = \hat{\Gamma}(l)$. This implies that $\hat{\gamma}$ will assign zero on $h(\cdot)$, if not, $\hat{\Gamma}(h) > \hat{\Gamma}(l)$, which is contradictory. $\square$

**Lemma B.1.** *For any function $w : Y \to \mathbb{R}$, reversely FOSD-ordered $\{\gamma(\cdot|\omega)\}_{\omega\in\Omega}$ and $\hat{\gamma}(\cdot)$ defined by (3.4), the following equality holds for all $\omega \in \Omega$:*

$$\int_{y\in Y} w(y)d\Gamma(y|\omega) = \int_{h\in\mathcal{D}} w(h(\omega))d\hat{\Gamma}(h). \tag{B.5}$$

*Proof.* Fix $\omega_0 \in \Omega$ and consider a function $h_{y_0,\omega_0}$ defined as follows:

$$h_{y_0,\omega_0} = \begin{cases} \bar{y}, & \text{if } \omega < \omega_0, \\ y_0, & \text{if } \omega \geq \omega_0, \end{cases} \tag{B.6}$$

where $\bar{y} = \max Y$.

From (3.4) and reversely FOSD-order of $\Gamma(\cdot|\omega)$, note that

$$\hat{\Gamma}(h_{y_0,\omega_0}) = \Gamma(h(\omega_0)|\omega_0) = \Gamma(y_0|\omega_0). \tag{B.7}$$

Define a set of functions $\mathcal{D}(y_0, \omega_0) \subset Y^\Omega$ as follows:

$$\mathcal{D}(y_0,\omega_0) = \left\{ h : \Omega \to Y \,\middle|\, \begin{array}{l} h: \text{ nonincreasing,} \\ h(\omega_0) = y_0 \end{array} \right\}. \tag{B.8}$$

Note that for all $y \leq y_0$ and $h \in \mathcal{D}(y, \omega_0)$, $h(\omega) \leq h_{y_0,\omega_0}(\omega)$. Since $\hat{\gamma}$ assigns zero to functions which are not nonincreasing,

$$\hat{\Gamma}(h_{y_0,\omega_0}) = \int_{y\leq y_0} \int_{h\in\mathcal{D}(y,\omega_0)} d\hat{\Gamma}(h)dy. \tag{B.9}$$

From total differentiations of (B.7) and (B.9) in terms of $y_0$, we can get

$$d\hat{\Gamma}(h_{y_0,\omega_0}) = d\Gamma(y_0|\omega_0) = \int_{h\in\mathcal{D}(y_0,\omega_0)} d\hat{\Gamma}(h)dy. \tag{B.10}$$

Then, the following equalities hold:

$$\begin{aligned}\int_{y\in Y} w(y)d\Gamma(y|\omega_0) &= \int_{y\in Y} w(y)\int_{h\in\mathcal{D}(y,\omega_0)} d\hat{\Gamma}(h)dy \\ &= \int_{y\in Y}\int_{h\in\mathcal{D}(y,\omega_0)} w(h(\omega_0))d\hat{\Gamma}(h)dy \\ &= \int_{h\in\mathcal{D}} w(h(\omega_0))d\hat{\Gamma}(h),\end{aligned}$$

the last equality is due to $\mathcal{D} = \bigsqcup_{y\in Y}\mathcal{D}(y,\omega_0)$. □

*Proof for Lemma 3.2.* Set $w(\cdot) = u(a_g(\cdot),\omega)$ in Lemma B.1. □

*Proof of Theorem 1.* Let $a_g(y)$ $(a_f(x))$ be the optimal decision at the third stage (i.e., the solution of (3.1)) with a signal $y$ $(x)$ from the information structure $g$ $(f)$. Note that the following equalities and inequalities hold (detailed steps will be presented in the following paragraph):

$$\begin{aligned}&\int_{\omega\in\Omega}\left[\int_{y\in Y} u(a_g(y),\omega)d\Gamma(y|x,\omega)\right]d\Lambda_f^x \\ &= \int_{\omega\in\Omega}\left[\int_{h\in\mathcal{D}} u(a_g(h(\omega)),\omega)d\hat{\Gamma}(h|x)\right]d\Lambda_f^x \\ &= \int_{h\in\mathcal{D}}\left[\int_{\omega\in\Omega} u(a_g(h(\omega)),\omega)d\Lambda_f^x\right]d\hat{\Gamma}(h|x) \\ &\leq \int_{h\in\mathcal{D}}\left[\int_{\omega\in\Omega} u(\phi(h),\omega)d\Lambda_f^x\right]d\hat{\Gamma}(h|x) \\ &\qquad\qquad\qquad\qquad \text{for some } \phi:\mathcal{D}\to A \\ &\leq \int_{h\in\mathcal{D}}\left[\int_{\omega\in\Omega} u(a_f(x),\omega)d\Lambda_f^x\right]d\hat{\Gamma}(h|x) \\ &= \int_{\omega\in\Omega} u(a_f(x),\omega)d\Lambda_f^x.\end{aligned}$$

The first equality holds from Lemma 3.2 and the second equality comes from a change of integration order. By the monotonicity condition of $g$, for all $h\in\mathcal{D}$, $a_g(h(\cdot)):\Omega\to$

$A$ is a decreasing function in terms of $\geq_A$. Then, by well-ordered condition, there exists $\phi(h) \in A$ such that $u(\phi(h), \omega) \geq u(a_g(h(\omega)), \omega)$ so that $\int_{\omega\in\Omega} u(\phi(h), \omega) d\Lambda_f^x \geq \int_{\omega\in\Omega} u(a_g(h(\omega)), \omega) d\Lambda_f^x$, thus, the third inequality holds. Since $a_f(x)$ is the solution of (3.1), the fourth inequality holds. The last inequality holds due to $\int_{h\in\mathcal{D}} d\hat{\Gamma}(h|x) = 1$.

Then, we can derive the following result:

$$
\begin{aligned}
V(g; u, \lambda) &= \int_{y\in Y} \int_{\omega\in\Omega} u(a_g(y), \omega) g(y|\omega) \lambda(\omega) d\omega dy \\
&= \int_{y\in Y} \int_{\omega\in\Omega} \int_{x\in X} u(a_g(y), \omega) \gamma(y|x, \omega) f(x|\omega) \lambda(\omega) dx\, d\omega\, dy \\
&= \int_{x\in X} \int_{\omega\in\Omega} \left[ \int_{y\in Y} u(a_g(y), \omega) d\Gamma(y|x, \omega) \right] d\Lambda_f^x\, dF_\lambda \\
&\leq \int_{x\in X} \int_{\omega\in\Omega} u(a_f(x), \omega) d\Lambda_f^x\, dF_\lambda = V(f; u, \lambda).
\end{aligned}
$$

□

### B.2.1 Discrete Example of $\hat{\gamma}$

The goal of this subsection is to illustrate the change of probability measure technique in Subsection B.2 by using a discrete case. Recall that the results on Subsection B.2 imply that a reversely FOSD-ordered noise would be transformed to a probability measure on decreasing functions. The following definition is useful in describing a decreasing function in discrete setup.

**Definition.** When $|\Omega| = n$ and $|Y| = L$, a matrix,

$$
B = \begin{pmatrix} b_{11} & \cdots & b_{1L} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nL} \end{pmatrix},
$$

is called a **monotone base** if it satisfies the following properties:

1. for all $1 \leq i \leq n$, there exists $1 \leq l(i) \leq L$ such that $b_{il(i)} = 1$ and $b_{il} = 0$ for all $l \neq l(i)$,
2. $l(i) \geq l(j)$ for all $1 \leq i \leq j \leq n$.

Note that any decreasing function $h : \Omega \to Y$ corresponds to a monotone base. To illustrate Proposition 3.1, consider a reversely FOSD-ordered noise $\{\gamma(\cdot|\omega)\}_{\omega\in\Omega}$ given by

$$\gamma = \begin{pmatrix} 0 & .7 & .3 \\ .1 & .7 & .2 \\ .3 & .5 & .2 \end{pmatrix}.$$

Since the goal is to transform $\gamma$ to a probability measure on decreasing functions, I will show how to decompose $\gamma$ into convex combination of monotone bases. As a first step, consider the components of $\gamma$ satisfying following properties:

1. $\gamma(y_l|\omega_i) \neq 0$,
2. for all $i' < i$, $\gamma(y_l|\omega_{i'}) = 0$,
3. for all $l' < l$, $\gamma(y_{l'}|\omega_i) = 0$,

and denote these properties as $(\star)$. This means that there is no non-zero component on the left and upper side of these components. Then, in the case of $\gamma$, the components satisfying $(\star)$ are $\gamma(y_2|\omega_1) = 0.7$ and $\gamma(y_1|\omega_2) = 0.1$. The location of the components satisfying $(\star)$ uniquely defines a monotone base such that 1s are written in the same row of the component satisfying $(\star)$ from the component to right before the next component satisfying $(\star)$ . In this case, the monotone base is

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The next step is to take a minimum of the components satisfying $(\star)$. In this case, I should take 0.1 between 0.1 and 0.7. Then, multiply this minimum with the monotone base generated from the components satisfying $(\star)$ and subtract it from $\gamma$. Note that the property $(\star)$ and reversely FOSD-order of $\gamma$ guarantee that the obtained matrix after subtraction is still composed of non-negative components and the number of zero components has increased. By repeating similar subtractions, eventually, zero matrix will be obtained and I am able to write $\gamma$ as follows:

$$
\begin{aligned}
\gamma = \begin{pmatrix} 0 & .7 & .3 \\ .1 & .7 & .2 \\ .3 & .5 & .2 \end{pmatrix} = \ & 0.1 \times \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \ 0.2 \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\
& + \ 0.4 \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} + \ 0.1 \times \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\
& + \ 0.2 \times \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}
$$

In this example, we can see that a reversely FOSD-ordered noise $\gamma$ can be represented as a convex combination of monotone bases, which correspond to decreasing functions. By multiplying $1 \times L$ matrix to the above equation, the result of Lemma B.1 can also be illustrated.

## B.3 Omitted Details in Subsection 3.3

As a tool for establishing results in this subsection, I consider a set of alternatives $A = \{0, 1\}$ and utility function

$$
u(a, \omega) = \begin{cases} -a \cdot \kappa & \text{if } \omega < \hat{\omega} \\ a \cdot (1 - \kappa) & \text{if } \omega \geq \hat{\omega} \end{cases} \tag{B.11}
$$

for some $\hat{\omega} \in \Omega$ and $\kappa \in \mathbb{R}$.[24] Note that $A$ satisfies well-ordered condition with respect to $u$. Then, for a prior $\lambda \in \Delta(\Omega)$, define

$$\begin{aligned}
f_\lambda(x|\omega < \hat{\omega}) &\equiv \frac{1}{\Lambda(\hat{\omega})} \int_{\omega < \hat{\omega}} f(x|\omega)\lambda(\omega) d\omega, \\
f_\lambda(x|\omega \geq \hat{\omega}) &\equiv \frac{1}{1 - \Lambda(\hat{\omega})} \int_{\omega \geq \hat{\omega}} f(x|\omega)\lambda(\omega) d\omega, \\
g_\lambda(y|\omega < \hat{\omega}) &\equiv \frac{1}{\Lambda(\hat{\omega})} \int_{\omega < \hat{\omega}} g(y|\omega)\lambda(\omega) d\omega, \\
g_\lambda(y|\omega \geq \hat{\omega}) &\equiv \frac{1}{1 - \Lambda(\hat{\omega})} \int_{\omega \geq \hat{\omega}} g(y|\omega)\lambda(\omega) d\omega.
\end{aligned}$$

Note that the expected utility after receiving a signal $x$ from information structure $f$ would be

$$\int_{\omega \in \Omega} u(a, \omega) d\Lambda_f^x(\omega) = a \cdot (L_f(x) - \kappa)$$

where $L_f(x) \equiv \frac{f_\lambda(x|\omega \geq \hat{\omega}) \cdot (1 - \Lambda(\hat{\omega}))}{f_\lambda(x)}$, i.e., likelihood ratio between $\omega \geq \hat{\omega}$ and $\omega < \hat{\omega}$.[25] Then, the optimal choice of $a \in A$ is 1 if $L_f(x) \geq \kappa$ and 0 if $L_f(x) < \kappa$. For a formal representation, define a set

$$\begin{aligned}
\tilde{X}(\hat{x}, \hat{\omega}, \lambda) &\equiv \{x \in X \mid L_f(x) < L_f(\hat{x})\} \\
&= \left\{ x \in X \mid \frac{f_\lambda(x|\omega \geq \hat{\omega})}{f_\lambda(x|\omega < \hat{\omega})} < \frac{f_\lambda(\hat{x}|\omega \geq \hat{\omega})}{f_\lambda(\hat{x}|\omega < \hat{\omega})} \right\} \cdot
\end{aligned}$$

If $\kappa$ is defined to be equal to $L_f(\hat{x})$, the optimal choice for $x \in \tilde{X}(\hat{x}, \hat{\omega}, \lambda)$ is $a = 0$, while the optimal choice for $x \in \tilde{X}^c(\hat{x}, \hat{\omega}, \lambda)$ is $a = 1$.

Based on these observations, when $\kappa = L_f(\hat{x})$, the expected utility from informa-

[24] This choice of utility function and subsequent technique are inspired by Chi [2015].

[25] To be precise, the likelihood ratio is $\frac{1 - \Lambda_f^x(\hat{\omega})}{\Lambda_f^x(\hat{\omega})}$, but the orders of $L_f(x)$ and $\frac{1 - \Lambda_f^x(\hat{\omega})}{\Lambda_f^x(\hat{\omega})}$ are invariant, thus, it is eligible to interpret $L_f(x)$ as a likelihood ratio.

tion structure $f$ would be

$$
\begin{aligned}
V(f;u,\lambda)=&\int_{\tilde{X}^c(\hat{x},\hat{\omega},\lambda)}(L_f(x)-\kappa)f_\lambda(x)dx\\
=&(1-\Lambda(\hat{\omega}))(1-\kappa)\left(1-\int_{\tilde{X}(\hat{x},\hat{\omega},\lambda)}f_\lambda(x|\omega\geq\hat{\omega})dx\right)\\
&-\Lambda(\hat{\omega})\kappa\left(1-\int_{\tilde{X}(\hat{x},\hat{\omega},\lambda)}f_\lambda(x|\omega<\hat{\omega})dx\right).
\end{aligned}
$$

Similarly, the optimal choice after receiving a signal $y$ from $g$ depends on $L_g(y)$ and $\tilde{Y}(\hat{y},\hat{\omega},\lambda)$. Note that if $g$ satisfies the MLRP, $\tilde{Y}(\hat{y},\hat{\omega},\lambda)=\{y\in Y\mid y<\hat{y}\}$, thus, monotonicity condition holds. Since the optimal choice is better than choosing $a=0$ for $y<\hat{y}$ and $a=1$ for $y\geq\hat{y}$, the expected utility from information structure $g$ would be

$$
\begin{aligned}
V(g;u,\lambda)\geq&\int_{y\geq\hat{y}}(L_g(y)-\kappa)g_Y(y)dy\\
=&(1-\Lambda(\hat{\omega}))(1-\kappa)(1-G_\lambda(\hat{y}|\omega\geq\hat{\omega}))\\
&-\Lambda(\hat{\omega})\kappa(1-G_\lambda(\hat{y}|\omega<\hat{\omega})).
\end{aligned}
$$

Then, informativeness for monotone decision problems implies $V(f;u,\lambda)\geq V(g;u,\lambda)$ and the following can be derived:

$$
\begin{aligned}
&\frac{G_\lambda(\hat{y}|\omega\geq\hat{\omega})}{f_\lambda(\hat{x}|\omega\geq\hat{\omega})}-\int_{\tilde{X}(\hat{x},\hat{\omega},\lambda)}\frac{f_\lambda(x|\omega\geq\hat{\omega})}{f_\lambda(\hat{x}|\omega\geq\hat{\omega})}dx\\
\geq&\frac{G_\lambda(\hat{y}|\omega<\hat{\omega})}{f_\lambda(\hat{x}|\omega<\hat{\omega})}-\int_{\tilde{X}(\hat{x},\hat{\omega},\lambda)}\frac{f_\lambda(x|\omega<\hat{\omega})}{f_\lambda(\hat{x}|\omega<\hat{\omega})}dx.
\end{aligned}
\tag{B.12}
$$

*Proof for Proposition 3.2.* I begin by stating a useful lemma:

**Lemma B.2.** *If $|X|=2$, $g$ satisfies the strict MLRP, and $f$ and $g$ satisfy (B.12) for all $\hat{x}\in X$, $\hat{y}\in Y$, $\hat{\omega}\in\Omega$ and $\lambda\in\Delta(\Omega)$, then for all $\omega_i<\omega_j<\omega_k$, $f(\hat{x}|\omega_i)\leq f(\hat{x}|\omega_j)$ implies $f(\hat{x}|\omega_j)\leq f(\hat{x}|\omega_k)$.*

*Proof.* I prove by contradicting two other possible cases:

1. $f(\hat{x}|\omega_k)<f(\hat{x}|\omega_i)\leq f(\hat{x}|\omega_j)$

In this case, there exists $a \in (0, 1]$ such that

$$f(x|\omega_i) = a \cdot f(x|\omega_j) + (1 - a) \cdot f(x|\omega_k)$$

for all $x \in X$.

Set $\hat{\omega} = \omega_j$ and $\lambda$ as follows:

$$\lambda(\omega) = \begin{cases} 1/2 & \text{if } \omega = \omega_i, \\ a/2 & \text{if } \omega = \omega_j, \\ (1-a)/2 & \text{if } \omega = \omega_k, \\ 0 & \text{otherwise.} \end{cases}$$

Note that the prior is deliberately chosen to satisfy $f_\lambda(x|\omega < \hat{\omega}) = f_\lambda(x|\omega \geq \hat{\omega})$ for all $x \in X$ and $\tilde{X}(\hat{x}, \hat{\omega}, \lambda) = \emptyset$. Then, (B.12) implies that

$$G_\lambda(\hat{y}|\omega \geq \hat{\omega}) \geq G_\lambda(\hat{y}|\omega < \hat{\omega})$$

and it contradicts the strict MLRP of $g$.

2. $f(\hat{x}|\omega_i) \leq f(\hat{x}|\omega_k) < f(\hat{x}|\omega_j)$

   In this case, there exists $b \in [0, 1)$ such that

$$f(x|\omega_k) = b \cdot f(x|\omega_i) + (1 - b) \cdot f(x|\omega_j)$$

for all $x \in X$.

Set $\hat{\omega} = \omega_k$ and $\lambda$ as follows:

$$\lambda(\omega) = \begin{cases} b/2 & \text{if } \omega = \omega_i, \\ (1-b)/2 & \text{if } \omega = \omega_j, \\ 1/2 & \text{if } \omega = \omega_k, \\ 0 & \text{otherwise.} \end{cases}$$

By a similar argument as in the first case,

$$G_\lambda(\hat{y}|\omega \geq \hat{\omega}) \geq G_\lambda(\hat{y}|\omega < \hat{\omega})$$

and it contradicts the strict MLRP of $g$.

$\square$

Consider $\omega_i < \omega_j$. If $f(x|\omega_i) = f(x|\omega_j)$, by considering a prior $\lambda$ that assigns $1/2$ for $\omega_i$ and $\omega_j$ (and 0 for other states) and applying (B.12) as in Lemma B.2, it contradicts the strict MLRP of $g$. Since $|X| = 2$, let $f(x_1|\omega_i) > f(x_1|\omega_j)$ and $f(x_2|\omega_i) < f(x_2|\omega_j)$.

By Lemma B.2, observe following facts:

1. If $\omega_j < \omega$, $f(x_2|\omega_j) < f(x_2|\omega)$. This is the direct result of the lemma.

2. If $\omega_i < \omega < \omega_j$, $f(x_2|\omega_i) < f(x_2|\omega) < f(x_2|\omega_j)$. If the first inequality fails, applying the lemma on $\hat{x} = x_1$ gives contradiction to $f(x_2|\omega_i) < f(x_2|\omega_j)$. Given the first inequality holds, the second inequality follows from the lemma.

3. If $\omega_i < \omega$, $f(x_2|\omega) < f(x_2|\omega_i)$. If the inequality fails, apply the lemma on $\hat{x} = x_1$ as in the previous case.

Now suppose $f(x_2|\omega)$ is not an increasing function on $\omega$. Then, there exists $\omega < \omega'$ with $f(x_2|\omega) > f(x_2|\omega')$ (Recall that they cannot be equalized). By the above facts, $(\omega, \omega')$ should be in the same interval. Consider the three cases:

1. If $\omega_j < \omega < \omega'$, note that $f(x_2|\omega_j) < f(x_2|\omega)$. By applying the lemma, $f(x_2|\omega) < f(x_2|\omega')$ and it is a contradiction.

2. If $\omega_i < \omega < \omega' < \omega_j$, note that $f(x_2|\omega_i) < f(x_2|\omega)$. By applying the lemma, $f(x_2|\omega) < f(x_2|\omega')$ and it is a contradiction.

3. If $\omega < \omega' < \omega_i$, note that $f(x_2|\omega') < f(x_2|\omega_i)$ and $f(x_1|\omega') > f(x_1|\omega_i)$. Also note that $f(x_1|\omega) < f(x_1|\omega')$ from $f(x_2|\omega) > f(x_2|\omega')$. By applying the lemma on $\hat{x} = x_1$, we have $f(x_1|\omega') < f(x_1|\omega_1)$ and it is a contradiction.

Therefore, $f(x_2|\omega)$ is an increasing function on $\omega$ and it means $\frac{f(x_2|\omega)}{f(x_1|\omega)}$ is increasing in $\omega$ so that the strict MLRP holds for $f$. $\square$

*Proof for Theorem 4.* Let $Y = \{y_1, y_2\}$ and $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$. Then, $g$ is summarized by $g_i \equiv g(y_1|\omega_i)$ and the MLRP of $g$ implies that $g_1 \geq g_2 \geq \cdots \geq g_n$. Define $\mathcal{L}_i : [0,1]^X \to [0,1]$ as

$$\mathcal{L}_i(h) = \int_{x \in X} h(x) f(x|\omega_i) dx.$$

Define $H(\omega_i) \equiv \mathcal{L}_i^{-1}(g_i)$ and

$$e(H) \equiv \{h' : X \to [0,1] \mid \exists h \in H \text{ s.th. } h'(x) \geq h(x) \ \forall x \in X\}.$$

Note that $e(\emptyset) = \emptyset$. Define

$$\begin{aligned}
\mathcal{H}_1 \equiv& H(\omega_1), \\
\mathcal{H}_2 \equiv& H(\omega_2) \cap e(\mathcal{H}_1), \\
&\vdots \\
\mathcal{H}_n \equiv& H(\omega_3) \cap e(\mathcal{H}_{n-1}).
\end{aligned}$$

I state two useful lemmas:

**Lemma B.3.** *When $|Y| = 2$ and $|\Omega| = n$, $g$ is a monotone quasi-garbling of $f$ if and only if $\mathcal{H}_n \neq \emptyset$.*

*Proof.* ($\Rightarrow$) If $g$ is a monotone quasi-garbling of $f$ and $\gamma : X \times \Omega \to \Delta(Y)$ is the garbling probability, $\gamma(y_1| \cdot, \omega_i) \in \mathcal{H}_i$.

($\Leftarrow$) Pick $h_n \in \mathcal{H}_n = H(\omega_n) \cap e(\mathcal{H}_{n-1})$. Then, from $h_n \in e(\mathcal{H}_{n-1})$, there exists $h_{n-1} \in \mathcal{H}_{n-1}$ such that $h_n(x) \geq h_{n-1}(x)$ for all $x \in X$. Define $h_i$ inductively so that $\{h_i\}_{1 \leq i \leq n}$ satisfies $h_i \in H(\omega_i)$ and $h_{j+1}(x) \geq h_j(x)$ for all $x \in X$ and $1 \leq j \leq n-1$. $\square$

**Lemma B.4.** *When $\mathcal{H}_i \neq \emptyset$, $\mathcal{H}_{i+1} \neq \emptyset$ if and only if*

$$g_{i+1} \geq \min_{h \in \mathcal{H}_i} \mathcal{L}_{i+1}(h).$$

*Proof.* ($\Rightarrow$) Let $\hat{h} \in \mathcal{H}_{i+1}$. By definition of $\mathcal{H}_{i+1}$, $g_{i+1} = \mathcal{L}_{i+1}(h)$ and there exists $h' \in \mathcal{H}_i$ such that $\hat{h}(x) \geq h'(x)$ for all $x \in X$. Therefore,

$$g_{i+1} = \mathcal{L}_{i+1}(\hat{h}) \geq \mathcal{L}_{i+1}(h') \geq \min_{h \in \mathcal{H}_i} \mathcal{L}_{i+1}(h).$$

($\Leftarrow$) Define $\mathbf{1} \in [0,1]^X$ as $\mathbf{1}(x) = 1$ for all $x \in X$. Then,

$$\mathcal{L}_{i+1}(\mathbf{1}) \geq g_{i+1} \geq \min_{h \in \mathcal{H}_i} \mathcal{L}_{i+1}(h)$$

and there exists $\alpha \in [0,1]$ and $\tilde{h} \in \mathcal{H}_i$ such that $g_{i+1} = \alpha \cdot \mathcal{L}_{i+1}(\mathbf{1}) + (1-\alpha) \cdot \mathcal{L}_{i+1}(\tilde{h}) = \mathcal{L}_{i+1}(\alpha \cdot \mathbf{1} + (1-\alpha) \cdot \tilde{h})$. Note that $\alpha \cdot \mathbf{1} + (1-\alpha) \cdot \tilde{h} \in H(\omega_{i+1})$ and $(\alpha \cdot \mathbf{1} + (1-\alpha) \cdot \tilde{h})(x) \geq \tilde{h}(x)$ for all $x \in X$, thus, $\mathcal{H}_{i+1} \neq \emptyset$. □

By Lemma B.3 and B.4, it is enough to show that $g_{i+1} \geq \min_{h \in \mathcal{H}_i} \mathcal{L}_{i+1}(h)$ for $i = 1, 2$.

For $i = 1$, $\min_{h \in \mathcal{H}_i} \mathcal{L}_{i+1}(h)$ is equivalent to

$$\min_h \int_{x \in X} h(x) f(x|\omega_2) dx$$

subject to

$$\int_{x \in X} h(x) f(x|\omega_1) dx = g_1 \qquad (\ \eta\ )$$

$$h(x) \geq 0 \qquad (\ \underline{\mu}_x\ )$$

$$1 \geq h(x) \qquad (\ \overline{\mu}_x\ )$$

where $\eta$, $\underline{\mu}_x$, and $\overline{\mu}_x$ are Lagrange multipliers and note that $\underline{\mu}_x$, $\overline{\mu}_x \geq 0$ (but $\eta$ is not).

Then, the first order condition w.r.t. $h(x)$ is

$$\underline{\mu}_x - \overline{\mu}_x = f(x|\omega_1) \left( \frac{f(x|\omega_2)}{f(x|\omega_1)} - \eta \right).$$

If $\eta \leq 0$, the right hand side of the equation is positive due to the full support condition on $f$. Since only one of $\underline{\mu}_x$ and $\overline{\mu}_x$ can be positive and the other is zero, $\underline{\mu}_x > 0$ and $\overline{\mu}_x = 0$ for all $x \in X$. It implies that $h(x) = 0$ for all $x \in X$ and cannot satisfy $\mathcal{L}_1(h) = g_1$. Therefore, $\eta$ has to be positive.

Then, the solution $h^*$ of the problem takes a form of

$$h^*(x) = \begin{cases} 1 & \text{if } \frac{f(x|\omega_2)}{f(x|\omega_1)} < \eta \\ k & \text{if } \frac{f(x|\omega_2)}{f(x|\omega_1)} = \eta \\ 0 & \text{if } \frac{f(x|\omega_2)}{f(x|\omega_1)} > \eta \end{cases}$$

where $k \in [0,1]$.

Set $\hat{x}$ to satisfy $\frac{f(\hat{x}|\omega_2)}{f(\hat{x}|\omega_1)} = \eta$, $\hat{y} = y_1$, $\hat{\omega} = \omega_2$ and $\lambda \in \Delta(\Omega)$ to put $1/2$ for $\omega_1$ and $\omega_2$ and 0 for $\omega_3$. Note that $f_\lambda(x|\omega < \hat{\omega}) = f(x|\omega_1)$, $f_\lambda(x|\omega \geq \hat{\omega}) = f(x|\omega_2)$ and

$$\mathcal{L}_j(h^*) = \int_{\tilde{X}(\hat{x},\hat{\omega},\lambda)} f(x|\omega_j)dx + k \cdot \int_{\dot{X}(\hat{x},\hat{\omega},\lambda)} f(x|\omega_j)dx \tag{B.13}$$

for $j = 1, 2$, where $\dot{X}(\hat{x},\hat{\omega},\lambda) = \{x \in X \mid \frac{f(x|\omega<\hat{\omega})}{f(\hat{x}|\omega<\hat{\omega})} = \frac{f(x|\omega\geq\hat{\omega})}{f(\hat{x}|\omega\geq\hat{\omega})}\}$.

By the inequality (B.12), the equation (B.13) and $g_1 = \mathcal{L}_1(h^*)$,

$$\begin{aligned}
&\frac{1}{f(\hat{x}|\omega_2)}\left(g_2 - \mathcal{L}_2(h^*)\right) \\
=&\frac{1}{f(\hat{x}|\omega_2)}\left(g_2 - \int_{\tilde{X}(\hat{x},\hat{\omega},\lambda)} f(x|\omega_2)dx - k \cdot \int_{\dot{X}(\hat{x},\hat{\omega},\lambda)} f(x|\omega_2)dx\right) \\
\geq&\frac{1}{f(\hat{x}|\omega_1)}\left(g_1 - \int_{\tilde{X}(\hat{x},\hat{\omega},\lambda)} f(x|\omega_1)dx - k \cdot \int_{\dot{X}(\hat{x},\hat{\omega},\lambda)} f(x|\omega_1)dx\right) \\
=&\frac{1}{f(\hat{x}|\omega_1)}\left(g_1 - \mathcal{L}_1(h^*)\right) = 0
\end{aligned}$$

Therefore, $g_2 \geq \mathcal{L}_2(h^*) = \min_{h\in\mathcal{H}_1} \mathcal{L}_2(h)$.

For $i = 2$, $\min_{h\in\mathcal{H}_i} \mathcal{L}_{i+1}(h)$ is equivalent to

$$\min_h \int_{x\in X} h(x) f(x|\omega_3)dx$$

subject to

$$\int_{x\in X} h(x) f(x|\omega_2)dx = g_2 \tag{$\eta_2$}$$

$$\int_{x\in X} h(x) f(x|\omega_1)dx \geq g_1 \tag{$\eta_1$}$$

$$h(x) \geq 0 \tag{$\underline{\mu}_x$}$$

$$1 \geq h(x) \tag{$\overline{\mu}_x$}$$

where $\eta_1$, $\eta_2$, $\underline{\mu}_x$, and $\overline{\mu}_x$ are Lagrange multipliers and note that $\eta_1$, $\underline{\mu}_x$, $\overline{\mu}_x \geq 0$ (but $\eta_2$ is not). Then, the first order condition w.r.t. $h(x)$ is

$$\underline{\mu}_x - \overline{\mu}_x = f(x|\omega_3) - \eta_2 f(x|\omega_2) - \eta_1 f(x|\omega_1). \tag{B.14}$$

When $\eta_2 \geq 0$, define $\hat{\omega} = \omega_3$ and a prior $\lambda$ as follows:

$$\lambda = \begin{cases} \frac{\eta_1}{\eta_1+\eta_2+1} & \text{if } \omega = \omega_1 \\ \frac{\eta_2}{\eta_1+\eta_2+1} & \text{if } \omega = \omega_2 \\ \frac{1}{\eta_1+\eta_2+1} & \text{if } \omega = \omega_3 \end{cases}$$

By using $\hat{\omega}$ and $\lambda$, (B.14) can be rewritten as

$$\frac{\underline{\mu}_x - \overline{\mu}_x}{f_\lambda(x|\omega < \hat{\omega})} = \frac{f_\lambda(x|\omega \geq \hat{\omega})}{f_\lambda(x|\omega < \hat{\omega})} - (\eta_1 + \eta_2).$$

Then, the solution $h^*$ of the problem takes a form of

$$h^*(x) = \begin{cases} 1 & \text{if } \frac{f_\lambda(x|\omega \geq \hat{\omega})}{f_\lambda(x|\omega < \hat{\omega})} < \eta_1 + \eta_2 \\ k & \text{if } \frac{f_\lambda(x|\omega \geq \hat{\omega})}{f_\lambda(x|\omega < \hat{\omega})} = \eta_1 + \eta_2 \\ 0 & \text{if } \frac{f_\lambda(x|\omega \geq \hat{\omega})}{f_\lambda(x|\omega < \hat{\omega})} > \eta_1 + \eta_2 \end{cases}$$

where $k \in [0, 1]$.

Set $\hat{x}$ to satisfy $\frac{f_\lambda(\hat{x}|\omega \geq \hat{\omega})}{f_\lambda(\hat{x}|\omega < \hat{\omega})} = \eta_1 + \eta_2$ and $\hat{y} = y_1$. Then, $g_2 = \mathcal{L}_2(h^*)$. By using the inequality (B.12) as in $i = 1$, it can be shown that $g_3 \geq \mathcal{L}_3(h^*) = \min_{h \in \mathcal{H}_2} \mathcal{L}_2(h)$.

When $\eta_2 < 0$, set $\hat{\omega} = \omega_2$ and a prior $\lambda$ as follows:

$$\lambda = \begin{cases} \frac{\eta_1}{\eta_1-\eta_2+1} & \text{if } \omega = \omega_1 \\ \frac{-\eta_2}{\eta_1-\eta_2+1} & \text{if } \omega = \omega_2 \\ \frac{1}{\eta_1-\eta_2+1} & \text{if } \omega = \omega_3 \end{cases}$$

Then, by going through the similar process, it can also be shown that $g_3 \geq \mathcal{L}_3(h^*) = \min_{h \in \mathcal{H}_2} \mathcal{L}_2(h)$. $\square$

## B.4 Omitted Details in Section 4.3

The following proposition characterizes the solution of the maximization problem (4.11).

**Proposition B.1** (Maskin and Riley [1984]). *When the posterior belief is $\Upsilon \in \Delta(\Omega)$, let the solution of (4.11) be $q_\Upsilon(\omega)$. Then, $q_\Upsilon(\omega)$ is characterized by a set of subinter-*

*vals*

$$\{[a_j, b_j] \subseteq \Omega = [0,1] \,|\, a_{j+1} > b_j\}_{j \in J}$$

*such that*

1. $\bar{q}(a_j) = \bar{q}(b_j)$ *for all* $j \in J$ *except for the case that* $a_j = 0$,

2. *for all* $a < b$ *with* $b \notin \cup_{j \in J}(a_j, b_j)$,

$$\int_a^b \frac{\partial I}{\partial q}(\bar{q}(b), s) d\Upsilon(s) \leq 0,$$

   *with equality if* $b = b_j$ *and* $a = a_j > 0$ *for some* $j \in J$,

3. *the solution* $q_\Upsilon(\omega)$ *is*

$$q_\Upsilon(\omega) = \begin{cases} \bar{q}(b_j) & \text{if } \omega \in [a_j, b_j] \text{ for some } j \in J, \\ \bar{q}(\omega) & \text{otherwise.} \end{cases}$$

*Proof.* See proposition 7 of Maskin and Riley [1984]. □

*Proof of Lemma 4.2.* Consider any $\{q(\omega; s), p(\omega; s)\}_{(\omega,s) \in \Omega \times \Omega} \in {A'_3}^\Omega$ such that for all $s' > s$

$$\{q(\cdot; s), p(\cdot; s)\} \geq_{A'_3} \{q(\cdot; s'), p(\cdot; s')\},$$

i.e., $q(\omega; s') \geq q(\omega; s)$ for all $s' > s$ and $\omega \in \Omega$. Then, consider $\{\hat{q}(\cdot), \hat{p}(\cdot)\}$ defined by

$$\hat{q}(\omega) = q(\omega; \omega), \quad \hat{p}(\omega) = \int_0^\omega v_q(\hat{q}(s), s) ds,$$

and will show that

1. $\{\hat{q}(\cdot), \hat{p}(\cdot)\} \in A'_3$,

2. $u(\hat{q}(\omega), \hat{p}(\omega), \omega) \geq u(q(\omega; \omega), p(\omega; \omega), \omega)$.

By the construction of $\hat{p}(\cdot)$, if $\hat{q}(\cdot)$ is nondecreasing, $\{\hat{q}(\cdot), \hat{p}(\cdot)\} \in A'_3$. For any $\omega' > \omega$,

$$\hat{q}(\omega') = q(\omega'; \omega') \geq q(\omega'; \omega) \geq q(\omega; \omega) = \hat{q}(\omega).$$

The first inequality is due to $\geq_{A_3'}$ order and the second inequality is due to the nondecreasingness of $q(\cdot;\omega)$, thus, $\{\hat{q}(\cdot), \hat{p}(\cdot)\} \in A_3'$.

From $u(q,p,\omega) = -c(q) + p$, note that

$$u(\hat{q}(\omega), \hat{p}(\omega), \omega) = -c(q(\omega;\omega)) + \int_0^\omega v_q(q(s,s),s)ds,$$
$$u(q(\omega;\omega), p(\omega;\omega), \omega) = -c(q(\omega;\omega)) + \int_0^\omega v_q(q(s,\omega),s)ds.$$

Then, from $q(s,s) \leq q(s,\omega)$ for all $s \leq \omega$ and $v_{qq} \leq 0$, the following inequality holds:

$$u(\hat{q}(\omega), \hat{p}(\omega), \omega) \geq u(q(\omega;\omega), p(\omega;\omega), \omega).$$

Therefore, $A_3'$ satisfies well-ordered condition with respect to $u$. □

*Proof of Lemma 4.2.* Consider any $\{q(\omega;s), p(\omega;s)\}_{(\omega,s)\in\Omega\times\Omega} \in {A_3'}^{\Omega}$ such that for all $s' > s$

$$\{q(\cdot;s), p(\cdot;s)\} \geq_{A_3'} \{q(\cdot;s'), p(\cdot;s')\},$$

i.e., $q(\omega;s') \geq q(\omega;s)$ for all $s' > s$ and $\omega \in \Omega$. Then, consider $\{\hat{q}(\cdot), \hat{p}(\cdot)\}$ defined by

$$\hat{q}(\omega) = q(\omega;\omega), \quad \hat{p}(\omega) = \int_0^\omega v_q(\hat{q}(s),s)ds,$$

and will show that

1. $\{\hat{q}(\cdot), \hat{p}(\cdot)\} \in A_3'$,
2. $u(\hat{q}(\omega), \hat{p}(\omega), \omega) \geq u(q(\omega;\omega), p(\omega;\omega), \omega)$.

By the construction of $\hat{p}(\cdot)$, if $\hat{q}(\cdot)$ is nondecreasing, $\{\hat{q}(\cdot), \hat{p}(\cdot)\} \in A_3'$. For any $\omega' > \omega$,

$$\hat{q}(\omega') = q(\omega';\omega') \geq q(\omega';\omega) \geq q(\omega;\omega) = \hat{q}(\omega).$$

The first inequality is due to $\geq_{A_3'}$ order and the second inequality is due to the nondecreasingness of $q(\cdot;\omega)$. Therefore, we get $\{\hat{q}(\cdot), \hat{p}(\cdot)\} \in A_3'$.

From $u(q,p,\omega) = -c(q) + p$, note that

$$u(\hat{q}(\omega), \hat{p}(\omega), \omega) = -c(q(\omega;\omega)) + \int_0^\omega v_q(q(s,s),s)ds,$$
$$u(q(\omega;\omega), p(\omega;\omega), \omega) = -c(q(\omega;\omega)) + \int_0^\omega v_q(q(s,\omega),s)ds.$$

Then, from $q(s,s) \leq q(s,\omega)$ for all $s \leq \omega$ and $v_{qq} \leq 0$, we get $u(\hat{q}(\omega), \hat{p}(\omega), \omega) \geq u(q(\omega;\omega), p(\omega;\omega), \omega)$. Therefore, $A_3'$ satisfies well-ordered condition with respect to $u$. □

*Proof of Lemma 4.3.* First, I show that the second best solution is monotone.

**Lemma B.5.** *For all $x' > x$ and $\omega \in \Omega$,*

$$\bar{q}_f(\omega; x) \geq \bar{q}_f(\omega; x'). \tag{B.15}$$

*Proof of Lemma B.5.* Define $H(q,\omega)$ as follows:

$$H(q,\omega) = \frac{v_q(q,\omega) - c'(q)}{v_{q\omega}(q,\omega)}. \tag{B.16}$$

Recall that $v_q - c' \geq 0$ for $q \leq q^*(\omega)$. Then, note that from the assumptions on $v$ and $c$,

$$H_q(q,\omega) = \frac{(v_{qq} - c'')v_{q\omega} - (v_q - c')v_{qq\omega}}{v_{q\omega}^2} \leq 0 \tag{B.17}$$

holds for $q \leq q^*(\omega)$.

From (4.13), we can see that

$$H(\bar{q}_f(\omega;x),\omega) = \frac{1 - \Lambda_f^x(\omega)}{\lambda_f^x(\omega)}, \quad H(\bar{q}_f(\omega;x'),\omega) = \frac{1 - \Lambda_f^{x'}(\omega)}{\lambda_f^{x'}(\omega)}. \tag{B.18}$$

Then, from the MLRP of $f$, we can get the MLRP of $\lambda_f$ and it implies

$$H(\bar{q}_f(\omega;x'),\omega) = \frac{1 - \Lambda_f^{x'}(\omega)}{\lambda_f^{x'}(\omega)} \geq \frac{1 - \Lambda_f^x(\omega)}{\lambda_f^x(\omega)} = H(\bar{q}_f(\omega;x),\omega). \tag{B.19}$$

Since $\bar{q}_f(\omega;x)$, $\bar{q}_f(\omega;x) \leq q^*(\omega)$, (B.15) holds from (B.17). □

Now, it is enough to show that for all $x' > x$ and $\omega \in \Omega$,

$$q_f(\omega; x) \geq q_f(\omega; x'). \tag{B.20}$$

Let the intervals characterizing the optimal solutions $q_f(\omega; x')$ and $q_f(\omega; x)$ be

$$\left\{[a'_{j'}, b'_{j'}]\right\}_{j' \in J'} \quad \text{and} \quad \{[a_j, b_j]\}_{j \in J}.$$

Assume the contrary of (B.20). Then,

$$E = \{\omega \in \Omega \,|q_f(\omega; x') > q_f(\omega; x)\} \subset \Omega \tag{B.21}$$

is nonempty. Note that

$$E \subset \Big(\bigcup_{j' \in J'} [a'_{j'}, b'_{j'}]\Big) \bigcup \Big(\bigcup_{j \in J} [a_j, b_j]\Big) \tag{B.22}$$

because both optimal solutions coincide with the second best solutions outside of the set on the right hand side and the second best solutions follows the inequality (B.15) so that they cannot be in $E$.

**Lemma B.6.** *If $E$ is nonempty, there exists $j' \in J'$ such that $a'_{j'} \in E$. Then,*

$$q_f(a'_{j'}; x) < q_f(a'_{j'}; x') = \bar{q}_f(a'_{j'}; x') \leq \bar{q}_f(a'_{j'}; x) \tag{B.23}$$

*holds so that there exists $j \in J$ such that $a'_{j'} \in (a_j, b_j]$.*

*Proof.* First, consider a case that there exists $\omega_0 \in E$ such that $\omega_0 \in [a'_{j'}, b'_{j'}]$ for some $j' \in J'$. Then, the following holds:

$$q_f(a'_{j'}; x) \leq q_f(\omega_0; x) < q_f(\omega_0; x') = q_f(a'_{j'}; x). \tag{B.24}$$

The first inequality holds from the nondecreasingness of $q_f(\cdot; x)$. The second inequality is from $\omega_0 \in E$. The third equality is from $\omega_0 \in [a'_{j'}, b'_{j'}]$. Therefore, $a'_{j'} \in E$.

Next, consider a case that there exists $\omega_0 \in E$ such that $\omega_0 \in [a_j, b_j]$ for some $j \in J$. Then,

$$\bar{q}_f(b_j; x') \leq \bar{q}_f(b_j; x) = q_f(\omega_0; x) < q_f(\omega_0; x') \leq q_f(b_j; x'). \tag{B.25}$$

The first inequality is by (B.15). The second equality is from $\omega_0 \in [a_j, b_j]$ and the third equality is from $\omega_0 \in E$. The last inequality is from the nondecreasingness of $q_f(\cdot|x')$. Then, we get $\bar{q}_f(b_j; x') < q_f(b_j; x')$, which implies that $b_j \in [a'_{j'}, b'_{j'}]$ for some $j' \in J'$. Also note that $q_f(b_j; x) = \bar{q}_f(b_j; x) < q_f(b_j; x')$, therefore $b_j \in E$. Then, it goes back to the first case that there exists $\omega \in E$ such that $\omega_0 \in [a'_{j'}, b'_{j'}]$ for some $j' \in J'$ by setting $\omega = b_j$.

Since an element in $E$ is either in $[a'_{j'}, b'_{j'}]$ for some $j' \in J'$ or $[a'_{j'}, b'_{j'}]$ for some $j' \in J'$ from (B.22), one of the above cases will be applied and there exists $j' \in J'$ such that $a'_{j'} \in E$. Then, we can see the first inequality of (B.23) holds from $a'_{j'} \in E$ and the second equality holds since $a'_{j'} \in [a'_{j'}, b'_{j'}]$. The last inequality is due to (B.15). From $q_f(a'_{j'}; x) < \bar{q}_f(a'_{j'}; x)$, $a'_{j'}$ should be in some $[a_j, b_j]$. Since $q_f(a'_{j'}; x) \neq \bar{q}_f(a'_{j'}; x)$, $a'_{j'}$ cannot be equal to $a_j$. □

For $a'_{j'} \in [a_j, b_j]$ satisfying Lemma B.6, let $\bar{q} = \bar{q}(b_j; x) = q_f(b_j; x) = q_f(a'_{j'}; x)$. Note that the following inequality holds

$$\bar{q} = q_f(s; x) = q_f(a'_{j'}; x) < q_f(a'_{j'}; x') \leq q_f(s; x'), \tag{B.26}$$

for all $s \in [a'_{j'}, b_j]$.

Based on (4.12), define $I^x$ and $I^{x'}$ as follows:

$$I^x(q, \omega) = -c(q) + v(q, \omega) - v_\omega(q, \omega) \cdot \frac{1 - \Lambda^x_f(\omega)}{\lambda^x_f(\omega)}, \tag{B.27}$$

$$I^{x'}(q, \omega) = -c(q) + v(q, \omega) - v_\omega(q, \omega) \cdot \frac{1 - \Lambda^{x'}_f(\omega)}{\lambda^{x'}_f(\omega)}. \tag{B.28}$$

Note that

$$0 > \int_{a'_{j'}}^{b_j} \frac{\partial I^x}{\partial q} (\bar{q}, s)\, d\Lambda^x_f(s), \tag{B.29}$$

from the second statement of Proposition B.1. (The equality holds only if the lowest point is equal to $a_j$)

For all $a < b$, note that

$$\begin{aligned}
&\frac{1}{\Lambda_f^x(b)-\Lambda_f^x(a)}\cdot\left(\int_a^b \frac{\partial I^x}{\partial q}(q,s)d\Lambda_f^x(s)\right)\\
&=\frac{1}{\Lambda_f^x(b)-\Lambda_f^x(a)}\cdot\left(\int_a^b \frac{\partial}{\partial s}\left[(c'(q)-v_q(q,s))\cdot(1-\Lambda_f^x(s))\right]ds\right)\\
&=\frac{(c'(q)-v_q(q,b))(1-\Lambda_f^x(b))-(c'(q)-v_q(q,a))(1-\Lambda_f^x(a))}{\Lambda_f^x(b)-\Lambda_f^x(a)}\\
&=v_q(q,a)-c'(q)+(v_q(q,a)-v_q(q,b))\frac{1-\Lambda_f^x(b)}{\Lambda_f^x(b)-\Lambda_f^x(a)}.
\end{aligned}$$

From the MLRP of $f$, we have

$$\frac{1-\Lambda_f^x(b)}{\Lambda_f^x(b)-\Lambda_f^x(a)}\le\frac{1-\Lambda_f^{x'}(b)}{\Lambda_f^{x'}(b)-\Lambda_f^{x'}(a)},$$

for $x' > x$ and $b > a$. Therefore, we can derive that

$$\int_a^b \frac{\partial I^x}{\partial q}(q,s)d\Lambda_f^x(s)\ge\frac{\Lambda_f^x(b)-\Lambda_f^x(a)}{\Lambda_f^{x'}(b)-\Lambda_f^{x'}(a)}\cdot\int_a^b\frac{\partial I^{x'}}{\partial q}(q,s)d\Lambda_f^{x'}(s). \tag{B.30}$$

Then, from (B.29) and (B.30), we get

$$0>\int_{a'_{j'}}^{b_j}\frac{\partial I^{x'}}{\partial q}(\bar{q},s)d\Lambda_f^{x'}(s). \tag{B.31}$$

Note that $\frac{\partial^2 I^{x'}}{\partial q^2}\le 0$ from $c''\ge 0$, $v_{qq}\le 0$, and $v_{qq\omega}\ge 0$. Then, together with (B.26), we get

$$0>\int_{a'_{j'}}^{b_j}\frac{\partial I^{x'}}{\partial q}(q_f(s;x'),s)d\Lambda_f^{x'}(s). \tag{B.32}$$

Let $[a'_{j'}, b'_{j'}]$, $\cdots$ , $[a'_{j'+k}, b'_{j'+k}]$ be all the subintervals intersecting with $[a'_{j'}, b_j]$. Note that $q_f(s;x')$ coincides with the second best solution outside those subintervals and

$\frac{\partial I^{x'}}{\partial q}(q_f(s,x'),s) = 0$ by the definition of the second best solution. Therefore, we get

$$\begin{aligned}
&\int_{a'_{j'}}^{b_j} \frac{\partial I^{x'}}{\partial q}(q_f(s;x'),s)d\Lambda_f^{x'}(s) \\
&= \int_{a'_{j'}}^{b'_{j'}} \frac{\partial I^{x'}}{\partial q}(\bar{q}_f(b'_{j'};x'),s)d\Lambda_f^{x'}(s) \\
&\quad + \cdots \\
&\quad + \int_{a'_{j'+k}}^{\min\{b_j,b'_{j'+k}\}} \frac{\partial I^{x'}}{\partial q}(\bar{q}_f(b'_{j'+k};x'),s)d\Lambda_f^{x'}(s).
\end{aligned} \tag{B.33}$$

Note that all the integrations are equal to zero except for the last integration by the second statement of Proposition B.1. When $b_j \geq b'_{j'+k}$, the last integration is equal to zero as well. When $b_j < b'_{j'+k}$, note that

$$\int_{b_j}^{b'_{j'+k}} \frac{\partial I^{x'}}{\partial q}(\bar{q}_f(b'_{j'+k};x'),s)d\Lambda_f^{x'}(s) < 0,$$

and

$$\int_{a'_{j'+k}}^{b'_{j'+k}} \frac{\partial I^{x'}}{\partial q}(\bar{q}_f(b'_{j'+k};x'),s)d\Lambda_f^{x'}(s) = 0,$$

thus,

$$\int_{a'_{j'+k}}^{b_j} \frac{\partial I^{x'}}{\partial q}(\bar{q}_f(b'_{j'+k};x'),s)d\Lambda_f^{x'}(s) > 0.$$

Then, from (B.33), we get

$$\int_{a'_{j'}}^{b_j} \frac{\partial I^{x'}}{\partial q}(q_f(s;x'),s)d\Lambda_f^{x'}(s) \geq 0, \tag{B.34}$$

which is contradiction to (B.32). Therefore, the assumption that $E$ is nonempty is not valid and it implies that $q_f(\omega;x) \geq q_f(\omega;x')$ for all $x' > x$ and $\omega \in \Omega$. $\square$

# References


Susan Athey. Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, 117(1):187–223, 2002.

Susan Athey and Jonathan Levin. The value of information in monotone decision problems. *Research in Economics*, 2017.

David Blackwell. Comparison of experiments. In *Second Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 93–102, 1951.

David Blackwell. Equivalent comparisons of experiments. *The annals of mathematical statistics*, 24(2):265–272, 1953.

David Harold Blackwell and MA Girshick. Theory of games and statistical decision. Technical report, 1954.

Antonio Cabrales, Olivier Gossner, and Roberto Serrano. Entropy and the value of information for investors. *The American economic review*, 103(1):360–377, 2013.

Chang Koo Chi. The value of information and dispersion. 2015.

Juan-José Ganuza and Jose S Penalva. Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica*, 78(3):1007–1030, 2010.

Roger Guesnerie and Jean-Jacques Laffont. A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm. *Journal of Public Economics*, 25(3):329–369, December 1984.

Ian Jewitt. Information order in decision and agency problems. Nuffield College, 2007.

Samuel Karlin and Herman Rubin. The theory of decision procedures for distributions with monotone likelihood ratio. *The Annals of Mathematical Statistics*, pages 272–299, 1956.

Son Ku Kim. Efficiency of an Information System in an Agency Model. *Econometrica*, 63(1):89–102, January 1995.

EL Lehmann. Comparing location experiments. *The Annals of Statistics*, pages 521–533, 1988.

Jonathan Levin. Information and the market for lemons. *RAND Journal of Economics*, 32(4):657–66, Winter 2001.

Eric Maskin and John Riley. Monopoly with incomplete information. *RAND Journal of Economics*, 15(2):171–196, Summer 1984.

Paul R. Milgrom. Good News and Bad News: Representation Theorems and Applications. *Bell Journal of Economics*, 12(2):380–391, Autumn 1981.

Michael Mussa and Sherwin Rosen. Monopoly and product quality. *Journal of Economic Theory*, 18(2):301–317, August 1978.

Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1): 58–73, 1981.

Marco Ottaviani and Andrea Prat. The Value of Public Information in Monopoly. *Econometrica, Econometric Society*, 69(6):1673–1683, November 2001.

Nicola Persico. Information acquisition in auctions. *Econometrica*, 68(1):135–148, 2000.

John K-H Quah and Bruno Strulovici. Comparative statics, informativeness, and the interval dominance order. *Econometrica*, 77(6):1949–1992, 2009.

Anne-Katrin Roesler. Information disclosure in markets: Auctions, contests, and matching markets. Technical report, 2015.

Juuso Toikka. Ironing without control. *Journal of Economic Theory*, 146(6):2510–2526, 2011.