A Note on Comparing Information Structures using Posterior State Distributions

Christian Hermelingmeier*

October 30, 2007

Abstract

The paper studies the relationship between two approaches for ranking information structures: Blackwell's sufficiency criterion and two criteria based on the sensitivity of the posterior state distributions with respect to changes in the signal. For this purpose Blackwell's criterion is reformulated using posterior distribution functions. It is shown that only in the case of binary information structures the former concept implies the latter ones and in general there is no inclusion.

Keywords: Information, Information Structure, Sufficiency Criterion
JEL Classification: D8

---

[*] Financial support from the German Science Foundation (DFG) under grant GRK 1134/1 (International Research Training Group 'Economic Behavior and Interaction Models EBIM') is gratefully acknowledged. The author is thankful to Bernhard Eckwert, Manfred Nermuth, Andreas Szczutkowski and Andreas Eickhoff for helpful comments and suggestions. Correspondence to: Christian Hermelingmeier, Fakultät für Wirtschaftswissenschaften, Universität Bielefeld, Postfach 100 131, D-33501 Bielefeld, Germany, chermelingmeier@wiwi.uni-bielefeld.de.

1 Introduction

Information structures generating an observable signal correlated with an unobservable parameter or state of nature are widely used to study the role of information in economic models. However, the question of ordering different information structures is quite delicate and various concepts of informativeness have been developed (e.g. Blackwell (1951), Lehmann (1988) and Kim (1995)).

In his seminal work Blackwell links the ordering of information structures to a statistical sufficiency criterion. An information structure P is said to be more informative than a structure Q if Q can be generated from P via a stochastic transformation. This means, observing a signal from Q can be seen as a result of a garbled transmission of a signal of P. This criterion is very restrictive but allows for the strong equivalence that every expected utility maximizer will prefer a more informative structure and vice versa. It can also be closely linked to the inequality of income distributions and the riskiness of lotteries as Nermuth (1992) points out. Jewitt (1997) shows the conditions under which the other concepts mentioned above are subsequently weaker than Blackwell's.

Recently, Eckwert and Zilcha (2004) approach the ordering of information structures comparing the posterior state distributions. They regard an information structure as more informative if the posterior distribution is more sensitive with respect to changes in the signal. This gives a very tractable criterion for a number of circumstances but is restricted to those information structures possessing the monotone likelihood ratio property. Strictly weaker in this context is the notion of informativeness used by Ganuza and Penalva (2006) which compares the distribution of the posterior distribution's mean in terms of dispersion.

This paper focuses on the relationship between the criteria of Blackwell (1951) and Eckwert and Zilcha (2004) as well as Ganuza and Penalva (2006) and clarifies it. For this purpose it is necessary to reformulate Blackwell's notion using posterior distribution functions. Using an equivalent formulation by Marschak and Miyasawa (1968), it turns out that if an information structure is more informative in Blackwell's sense then every posterior distribution function is a convex combination of the posterior distribution functions generated by the alternative structure. Although it is argued that the sufficiency criterion implies a kind of higher dispersion of the posterior state distribution with respect to the signal, it is shown that only in the special case of binary information structures, i.e. two states and two signals, the first concept implies the other ones which are in this case equivalent. Beyond that, no more general inclusion concerning the relationship of the criteria holds. This highlights that relating results derived by the use of the two different approaches seems to be appropriate only in the binary case.

The paper is organized as follows. In section 2 Blackwell's criterion and the criteria of Eckwert and Zilcha and Ganuza and Penalva are introduced. In section 3 their relationship is analyzed. Concluding remarks are given in section 4. Some technical results are relegated to a separate appendix.

2 Comparing Information Structures

Before Blackwell's criterion is presented and discussed, the used notion of an information structure is shortly introduced. This is complemented by the generation of the posterior state distributions using Bayes's theorem.

The underlying uncertain situation is as follows. Tomorrow, nature will choose one out of n possible states denoted by $\omega_i \in \mathbb{R}, i = 1, ..., n$. A priori beliefs of tomorrow's state of nature are represented by probabilities $\pi_i > 0$ with $\sum_{i=1}^{n} \pi_i = 1$ and $\pi = (\pi_1, ..., \pi_n)$.

Now, assume that there are two alternative information structures P and Q, represented by row-stochastic[1] matrices $[P_{ij}]_{n \times m_P}$ and $[Q_{ik}]_{n \times m_Q}$. The signals of P are denoted by $y_j \in \mathbb{R}, j = 1, ..., m_P$. P_{ij} is the conditional probability of receiving signal y_j given state ω_i. Corresponding notation is used for Q. The a priori occurrence probabilities for the signals are

$$p_j^* = \sum_i \pi_i P_{ij} \quad \text{and} \quad q_k^* = \sum_i \pi_i Q_{ik} \tag{1}$$

with notation $p^* = (p_1^*, ..., p_{m_P}^*)$ and $q^* = (q_1^*, ..., q_{m_Q}^*)$. It is assumed that $p_j^* > 0$ and $q_k^* > 0$ for $j = 1, ..., m_P$ and $k = 1, ..., m_Q$ since any signal occur-

---

[1] A matrix is row-stochastic if all entries are non-negative and sum to one in each row.

ring with zero probability can obviously be removed from the structure. After observing a signal from an information structure, the a priori probability beliefs can be revised using Bayes's theorem and the a posteriori probability of state ω_i given signal y_j received from structure P is denoted by P^*_{ji}. With analogous definitions for Q, one gets two row-stochastic matrices $[P^*_{ji}]_{m_P \times n}$ and $[Q^*_{ki}]_{m_Q \times n}$ given by

$$P^*_{ji} = \frac{\pi_i P_{ij}}{p^*_j} \quad \text{and} \quad Q^*_{ki} = \frac{\pi_i Q_{ik}}{q^*_k}. \tag{2}$$

These are well defined since every signal will occur with at least some positive probability by assumption.

2.1 The Criterion of Blackwell

In order to compare information structures Blackwell (1951) formalizes the following intuitive idea. If the transmission of a signal generated by an information structure is garbled by a stochastic transformation uncorrelated with the true state of nature it loses informativeness. Therefore he regards an information structure as more informative than an alternative structure if the latter can be generated from the former by adding some random transmission error uncorrelated with the true state of nature.

Definition 1 *(Blackwell 1951) Information structure P is more informative than Q, $P \succeq^i Q$, if and only if there exists a row-stochastic matrix $\Gamma = [\gamma_{jk}]_{m_P \times m_Q}$, such that*

$$Q = P\Gamma. \tag{3}$$

That is $Q_{ik} = \sum_j P_{ij}\gamma_{jk}$, so that γ_{jk} can be interpreted as the conditional probability that when signal y_k is received from Q signal y_j was actually sent by P. Obviously, a transformation matrix Γ does not exist for every pair of information structures such that (3) holds. Therefore the criterion imposes only a partial ordering on the set of information structures. Definition 1 can be equivalently stated using the posterior matrices as follows.

Proposition 1 *(Marschak and Miyasawa 1968) Information structure P is more informative than Q if and only if there exists a row-stochastic matrix*

$\Gamma^* = [\gamma^*_{kj}]_{m_Q \times m_P}$, *such that the two following conditions are satisfied:*

$$(i) \quad \Gamma^* P^* = Q^* \tag{4}$$

$$(ii) \quad p^* = q^* \Gamma^* \tag{5}$$

Let P^*_j and Q^*_k denote the jth and kth rows of P^* and Q^*, i.e. the a posteriori distributions conditional on y_j and y_k, respectively. Condition (i) can then be written as

$$Q^*_k = \sum_j \gamma^*_{kj} P^*_j \tag{6}$$

for all $k = 1, ..., m_Q$. This means that the a posteriori distribution induced by observing signal y_k from information structure Q is a convex combination of the a posteriori distributions induced by the signals of structure P. For an individual state of nature ω_i this gives

$$Q^*_{ki} = \sum_j \gamma^*_{kj} P^*_{ji}. \tag{7}$$

Condition (i) alone is not sufficient for $P \overset{\mathrm{i}}{\succeq} Q$ in general, although stated in Bielinska-Kwapisz (2003). Condition (ii) is needed to show that the constructed garbling matrix Γ in line with (3) is in fact row-stochastic. Moreover, this can be interpreted as condition (ii) assuring that the given posterior matrices are consistent with being generated from the same a priori distribution. To see this, note that the prior can be calculated from the posteriors and signal occurrence probabilities by $p^* P^*$ and $q^* Q^*$, respectively. Conditions (i) and (ii) then imply $p^* P^* = q^* \Gamma^* P^* = q^* Q^*$.

Condition (i) of proposition 1 can equivalently be stated using the conditional distribution functions generated by the two information structures. This will be useful for the comparison of the sufficiency criterion with the other criteria. It is assumed without loss of generality that $\omega_k > \omega_l$ if $k > l$. Let $F(\omega_i | y_j)$ then denote the distribution function conditional on signal y_j from P evaluated at ω_i, i.e. for all $i = 1, ..., n$ and $j = 1, ..., m_P$ it is given by

$$F(\omega_i | y_j) = \sum_{l=1}^{i} P^*_{jl}. \tag{8}$$

$G(\omega_i | y_k)$ is the analogous term for a signal y_k generated by structure Q.

Lemma 1 *There exists a row-stochastic matrix* $\Gamma^* = [\gamma^*_{kj}]_{m_Q \times m_P}$*, such that* $\Gamma^* P^* = Q^*$ *if and only if for all* $i = 1, ..., n$ *and* $k = 1, ..., m_Q$

$$G(\omega_i|y_k) = \sum_j \gamma^*_{kj} F(\omega_i|y_j). \tag{9}$$

Proof: Suppose $\Gamma^* P^* = Q^*$ for such a Γ^*. Thus from (8) follows

$$G(\omega_i|y_k) = \sum_{l \leq i} Q^*_{kl} = \sum_{l \leq i} \sum_j \gamma^*_{kj} P^*_{jl} = \sum_j \gamma^*_{kj} \sum_{l \leq i} P^*_{jl} = \sum_j \gamma^*_{kj} F(\omega_i|y_j). \tag{10}$$

Now, suppose (9) is true. For $i = 1$ one gets

$$Q^*_{k1} = G(\omega_1|y_k) = \sum_j \gamma^*_{kj} F(\omega_1|y_j) = \sum_j \gamma^*_{kj} P^*_{j1} \tag{11}$$

and for $1 < i \leq n$ it follows

$$Q^*_{ki} = G(\omega_i|y_k) - G(\omega_{i-1}|y_k) = \sum_j \gamma^*_{kj} \left[F(\omega_i|y_j) - F(\omega_{i-1}|y_j)\right] = \sum_j \gamma^*_{kj} P^*_{ji}. \tag{12}$$

Thus, (7) holds for all $i = 1, ..., n$ and $\Gamma^* P^* = Q^*$ is shown. □

Combined with proposition 1 the following equivalent formulation of the sufficiency criterion can now be made.

Proposition 2 *Information structure* P *is more informative than* Q *if and only if there exists a row-stochastic matrix* $\Gamma^* = [\gamma^*_{kj}]_{m_Q \times m_P}$ *such that the following two conditions are satisfied:*

$$(i) \quad \forall i = 1, ..., n: \quad G(\omega_i|y_k) = \sum_j \gamma^*_{kj} F(\omega_i|y_j) \tag{13}$$

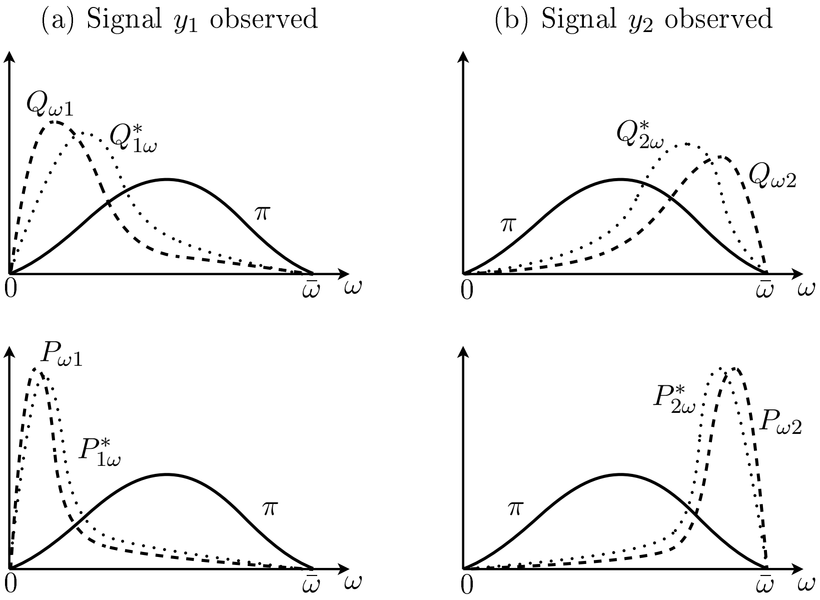
$$(ii) \quad p^* = q^* \Gamma^* \tag{14}$$

Proof: Immediate consequence of proposition 1 and lemma 1. □

This equivalent formulation of Blackwell's sufficiency criterion using posterior distribution functions seems not to be very well suited for application. To find a more tractable characterization of informativeness in terms of the posterior state distributions, a closer look on their generation by the use of Bayes's theorem in (2) seems to be an appropriate way to continue.

2.2 The Generation of State Posteriors

The Bayesian belief-revision process is illustrated in figure 1. For a convenient graphical representation a continuous rather than discrete state space is assumed, running from a lower limit $\omega = 0$ to an upper limit $\omega = \bar{\omega}$. In all diagrams π is the given a priori probability density function, depicted by the black line.


(a) Signal y_1 observed

(b) Signal y_2 observed

Figure 1: Bayesian revision of probability beliefs.

Now, consider an information structure Q and two alternative signals y_1 and y_2. As depicted by the dashed lines, the first signal has a much greater likelihood $Q_{\omega 1}$ if the true state of nature is relatively low (upper left diagram), while the likelihood $Q_{\omega 2}$ of the latter signal is greater for higher states (upper right diagram). Starting from the same a priori probability distribution in both panels, the differing likelihood curves lead to differing a posteriori probability density functions. As depicted in each panel, the dotted posterior distributions $Q^*_{1\omega}$ and $Q^*_{2\omega}$, respectively, are a kind of compromise or average of the two other curves.[2]

---

[2] For a detailed discussion see Hirshleifer and Riley (1992).

What influence does the informativeness of the underlying information structure have on this process? Consider an alternative information structure P which is more informative than structure Q. The greater mass of information is reflected in a 'tighter' likelihood curve as depicted in the lower left and right diagrams. Therefore, if the information contained in the observed signal becomes more informative, the a priori probability distribution has less influence on the posterior's shape and the state posteriors tend more towards the likelihoods. This means the observed signal has a greater impact on the posteriors, i.e. they are more 'dispersed' with respect to the signal. For example, consider the case of a totally uninformative and a fully informative information structure. In the first case, since the signal distribution does not depend on the true state of nature, the posterior distribution will not depend on the observed signal. In the second case, every signal reveals a certain state of nature to be the true one and the posterior distribution concentrates all probability at this state.
That this characterization of informativeness in terms of a higher dispersion with respect to the observed signal seems to be in line with Blackwell's notion is supported by the following fact which is an obvious implication of proposition 2.

Lemma 2 *If information structure P is more informative than Q then for every $i = 1, ..., n$*

$$\min_j F(\omega_i|y_j) \leq \min_k G(\omega_i|y_k) \leq \max_k G(\omega_i|y_k) \leq \max_j F(\omega_i|y_j). \qquad (15)$$

For any fixed state of nature ω_i the minimum and maximum of the posterior distribution function of the more informative information structure are more dispersed and therefore this posterior distribution covers a broader range of values. Obviously, this is only a weak implication. Hence, the question arises what a more precise characterization of Blackwell's sufficiency criterion in terms of the dispersion of the posterior state distributions with respect to the observed signal looks like and if this could be the criteria presented next.

2.3 Sensitivity Criteria

Eckwert and Zilcha (2004) use a ranking criterion for information structures based on the sensitivity of the posterior distribution function with respect to a change in the observed signal. A strictly weaker criterion comparing the dispersion of the posterior distribution's mean is proposed by Ganuza and Penalva (2006). Both criteria will be presented in the following. The question in which cases this kind of sensitivity is actually a necessary or sufficient condition for more informativeness in the sense of Blackwell is answered in section 3.

To have a suitable structure on the set of signals, attention is now restricted to information structures satisfying the monotone likelihood ratio property (MLRP): $y_s > y_t$ implies that for every nondegenerate a priori distribution the posterior distribution conditional on y_s dominates the posterior distribution conditional on y_t in the sense of first order stochastic dominance (Milgrom 1981). For convenience and without loss of generality it is assumed that $y_s > y_t$ if $s > t$. Furthermore, it assumed that $m_P = m_Q$, i.e. the information structures have the same number of potential signals. Without loss of generality the signal sets will therefore be taken to be identical.

Eckwert and Zilcha regard an information structure as more informative if the posterior state distribution is more sensitive with respect to changes in the signal. This means that the observable signal realizations have a uniformly stronger impact on the a posteriori distribution of states. Instead of using derivatives, a discrete version of their concept is given below.

Definition 2 *(Signal-Sensitivity) Information structure P is said to be more signal-sensitive than Q, denoted $P \succeq^{ss} Q$, if and only if for all ω_i and all $y_s > y_t$*

$$G(\omega_i|y_t) - G(\omega_i|y_s) \leq F(\omega_i|y_t) - F(\omega_i|y_s). \tag{16}$$

Ganuza and Penalva use the expected value of the posterior distributions for the comparison of information structures rather than the whole distributions. Therefore denote by $W^P(y_s)$ the mean of the posterior distribution conditional on signal y_s of information structure P and let $W^Q(y_s)$ be the analogous term for structure Q.

Since the generation of a signal can be seen as a random variable Y so can the expected value conditional on it. Intuitively, under an information structure with more informative signals the distribution of the updated expected value will become more spread out. For example, with a totally uninformative structure the distribution will be concentrated at the expected value and for a structure representing perfect information it will be the a priori distribution. Ganuza and Penalva regard an information structure as more informative if the distribution of $W^P(Y)$ is more Bickel-Lehmann disperse[3] than the distribution of $W^Q(Y)$. Using the notion of supermodularity the criterion is as follows.

Definition 3 *(Ganuza and Penalva 2006) The information structure P is more informative in terms of supermodular-precision than information structure Q, denoted $P \succeq^{sm} Q$, if and only if for all $y_s > y_t$*

$$W^P(y_s) - W^P(y_t) \geq W^Q(y_s) - W^Q(y_t). \tag{17}$$

The following result concerning the relationship of signal-sensitivity and supermodular-precision is proved in an earlier version of Ganuza and Penalva (2006). As already mentioned above, it states that the former concept is more restrictive than the latter one.

Proposition 3 *If information structure P is more signal-sensitive than Q then P is more informative in terms of supermodular-precision than Q.*

3 Relating the Criteria

The following results clarify the relation of Blackwell's sufficiency criterion and the criteria of signal-sensitivity and supermodular-precision. The case of binary information structures will turn out to be of special importance. The first result below is already shown in the appendix of Ganuza and Penalva (2006).

---

[3] A random variable X with cumulative distribution function F is more disperse in the sense of Bickel and Lehmann (1976) than another random variable Y with cumulative distribution function G if for all u, $v \in [0,1]$ with $u < v$

$$F^{-1}(v) - F^{-1}(u) \geq G^{-1}(v) - G^{-1}(u).$$

Proposition 4 *$P \succ^{i} Q$ does not necessarily imply $P \succ^{sm} Q$.*

Having in mind that, according to proposition 3, signal-sensitivity as a ranking criterion is more restrictive than supermodular-precision, this result allows to draw quickly a conclusion concerning the relationship of Blackwell's criterion and the sensitivity criterion.

Proposition 5 *$P \succ^{i} Q$ does not necessarily imply $P \succ^{ss} Q$.*

Proof: Suppose $P \succ^{\mathrm{i}} Q$ implies $P \succ^{\mathrm{ss}} Q$. Then, using lemma 3, $P \succ^{\mathrm{sm}} Q$ would follow from $P \succ^{\mathrm{i}} Q$. A contradiction to proposition 1. $\square$

Thus the concepts of signal-sensitivity and supermodular-precision do not characterize Blackwell's notion of informativeness in general. However, if attention is restricted to the case of binary information structures the situation changes. Below it will be shown that in this set of information structures $P \succ^{\mathrm{i}} Q$ in fact implies $P \succ^{\mathrm{ss}} Q$.

Note, that in the case of two states of nature and two signals the signal-sensitivity condition (16) is equivalent to just the following inequality

$$G(\omega_1|y_1) - G(\omega_1|y_2) \leq F(\omega_1|y_1) - F(\omega_1|y_2), \tag{18}$$

i.e. an information structure is more signal-sensitive if and only if a signal change has a greater impact on the posterior probability for state ω_1. This fact is used to prove the following result.

Proposition 6 *If P and Q are binary information structures then $P \succ^{i} Q$ implies $P \succ^{ss} Q$.*

Proof: Suppose $P \succ^{\mathrm{i}} Q$. If information structure Q is totally uninformative, the posteriors induced by its two signals are identical and therefore (18) is obviously satisfied since $G(\omega_1|y_1) - G(\omega_1|y_2)=0$.[4]

Now let Q not be a totally uninformative structure, i.e. its posteriors are not identical. Using corollary 2 and the fact that the matrix Γ^* is row-stochastic

---

[4] Note that due to the MLRP both sides in (18) are non-negative.

one gets

$$\begin{aligned} G(\omega_1|y_1) - G(\omega_1|y_2) &= \sum_{j=1}^{2} \left[\gamma^*_{1j} - \gamma^*_{2j}\right] F(\omega_1|y_j) \qquad (19) \\ &= \left[\gamma^*_{11} - \gamma^*_{21}\right] F(\omega_1|y_1) + \left[\gamma^*_{12} - \gamma^*_{22}\right] F(\omega_1|y_2) \\ &= \left[\gamma^*_{11} + \gamma^*_{22} - 1\right] F(\omega_1|y_1) - \left[\gamma^*_{11} + \gamma^*_{22} - 1\right] F(\omega_1|y_2) \\ &= \left[\gamma^*_{11} + \gamma^*_{22} - 1\right] \left[F(\omega_1|y_1) - F(\omega_1|y_2)\right] \end{aligned}$$

Due to the MLRP the posterior distributions are ordered in the sense of first order stochastic dominance, as already mentioned. In the binary case this is equivalent to the posterior matrices having trace greater than or equal to one. Now, the posteriors of Q are not identical by assumption and since $P \succ^{\mathrm{i}} Q$ the posteriors of P are not identical, too, which is an obvious implication of condition (i) in proposition 2. Therefore the inequality is strict, i.e. the posterior matrices have trace greater than one.
Since the stochastic matrices possessing this property form a group (see the appendix for details) and $Q^* = \Gamma^* P^*$ by Lemma 1 the transformation matrix Γ^* is of this type as well and also unique (Mac Lane and Birkhoff 1967, p. 77). Thus $0 < \gamma^*_{11} + \gamma^*_{22} - 1 \leq 1$ and this leads to

$$\left[\gamma^*_{11} + \gamma^*_{22} - 1\right] \left[F(\omega_1|y_1) - F(\omega_1|y_2)\right] \leq F(\omega_1|y_1) - F(\omega_1|y_2). \qquad (20)$$

Combine (19) and (20) to get

$$G(\omega_1|y_1) - G(\omega_1|y_2) \leq F(\omega_1|y_1) - F(\omega_1|y_2), \qquad (21)$$

i.e. (18) is satisfied and therefore $P \succ^{\mathrm{ss}} Q$ is shown. □

Proposition 6 shows that for the comparison of binary information structures the Blackwell criterion is more restrictive than the signal-sensitivity criterion. But this result is only true for this special case as is shown next.

Proposition 7 *The assumption of binary information structures is necessary for proposition 6 to hold.*

Proof: Consider a case of three states of nature and three signals.[5] Let the a priori distribution be given by

$$\pi = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right) \tag{22}$$

and the information structures be represented by the matrices

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} \frac{1}{2} & \frac{3}{4} & 0 \\ 0 & \frac{3}{4} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \tag{23}$$

Both possess the MLRP and $P \succ^{\mathrm{i}} Q$, because $Q = P\Gamma$ holds for the matrix

$$\Gamma = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \tag{24}$$

The signal occurrence probabilities are

$$p^* = \left(\frac{1}{8}, \frac{1}{4}, \frac{5}{8}\right) \quad \text{and} \quad q^* = \left(\frac{1}{16}, \frac{5}{8}, \frac{5}{16}\right) \tag{25}$$

and the posterior state distributions are given by

$$P^* = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{5} & \frac{4}{5} \end{pmatrix} \quad \text{and} \quad Q^* = \begin{pmatrix} 1 & 0 & 0 \\ \frac{3}{10} & \frac{3}{10} & \frac{2}{5} \\ 0 & \frac{1}{5} & \frac{4}{5} \end{pmatrix}. \tag{26}$$

The posterior state distributions react to a change in the signal as follows:

$$\begin{aligned} F(\omega_1|y_1) - F(\omega_1|y_2) = 0.5 < 0.7 = G(\omega_1|y_1) - G(\omega_1|y_2) \\ F(\omega_2|y_1) - F(\omega_2|y_2) = 0.0 < 0.4 = G(\omega_2|y_1) - G(\omega_2|y_2) \\ F(\omega_1|y_2) - F(\omega_1|y_3) = 0.5 > 0.3 = G(\omega_1|y_2) - G(\omega_1|y_3) \\ F(\omega_2|y_2) - F(\omega_2|y_3) = 0.8 > 0.4 = G(\omega_2|y_2) - G(\omega_2|y_3) \end{aligned} \tag{27}$$

---

[5] It suffices to show the claim for this case. A counterexample for any higher number of states and signals can easily be constructed by induction. Just add a row and a column to the matrices P and Q with one additional entry, all entries zero but the last one. This will not change (27). Repeat this a sufficient number of times.

Having definition 2 in mind, neither $P \not\succeq^{\mathrm{ss}} Q$ nor $Q \not\succeq^{\mathrm{ss}} P$ holds. $\square$

Propositions 6 and 7 state that in the binary case Blackwell's criterion is stronger than the signal-sensitivity criterion and that this implication holds only for the binary case. The relation of Blackwells criterion and the concept of supermodular-precision is exactly the same since regarding only binary information structures definitions 2 and 3 coincide.

Proposition 8 *In the binary case* $P \succeq^{ss} Q$ *if and only if* $P \succeq^{sm} Q$.

Proof: $P \succeq^{\mathrm{ss}} Q$ implies $P \succeq^{\mathrm{sm}} Q$ according to proposition 3 in general. Observing that due to $P \succeq^{\mathrm{ss}} Q$ being equivalent to (18), definition 2 generates a complete ordering on the set of binary information structures. Thus, in this case the two concepts have to be equivalent.[6] $\square$

Hence, propositions 6 and 7 in light of proposition 8 directly lead to the following relation of sufficiency and supermodular-precision.

Proposition 9 *Only in the binary case* $P \succeq^{i} Q$ *implies* $P \succeq^{sm} Q$.

Up to now it was stated that a more informative binary information structure is more informative in terms of signal-sensitivity as well as in terms of supermodular-precision. But, as shown by a counterexample below, a more signal-sensitive information structure need not be more informative in the sense of Blackwell, not even in the binary case.

Proposition 10 $P \succeq^{ss} Q$ *does not necessarily imply* $P \succeq^{i} Q$. *This holds even for the binary case.*

Proof: Consider the binary information structures given by[7]

$$P = \begin{pmatrix} 0.75 & 0.25 \\ 0.4 & 0.6 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{pmatrix}. \tag{28}$$

---

[6] Of course, the equivalence of conditions (16) and (17) can also be shown directly.

[7] It suffices to construct a counterexample for the binary case for the same reason as given in footnote 5.

The structures cannot be compared in the sense of Blackwell, i.e. neither $P \succ^{\mathrm{i}} Q$ nor $Q \succ^{\mathrm{i}} P$. This is easy to check using the criterion's characterization for the binary case (see lemma 3 in the appendix).
Now, set $\pi = (1/2, 1/2)$ for the a priori distribution.[8] Straightforward calculation leads to

$$G(\omega_1|y_1) - G(\omega_1|y_2) = \frac{2}{3} - \frac{4}{11} \leq \frac{15}{23} - \frac{5}{17} = F(\omega_1|y_1) - F(\omega_1|y_2). \qquad (29)$$

Thus, condition (18) holds and so P is more signal-sensitive than Q. $\square$

Using again the equivalence of signal-sensitivity and supermodular-precision in the binary case, proposition 10 directly gives the following.

Proposition 11 *$P \succ^{sm} Q$ does not necessarily imply $P \succ^{i} Q$. This holds even for the binary case.*

The set of pairs of binary information structures comparable in the sense of Blackwell is therefore a proper subset of the set of pairs comparable by the equivalent criteria of signal-sensitivity and supermodular-precision and the ordering is preserved. That means, in the binary case the Blackwell criterion is strictly stronger than the two other criteria. This relation of the three concepts is summarized in the following.

Corollary 1 *If P and Q are binary information structures the criteria are related as follows:*

$$P \succ^{i} Q \ \overset{\Rightarrow}{\not\Leftarrow}\ P \succ^{ss} Q \ \Leftrightarrow\ P \succ^{sm} Q \qquad (30)$$

4 Concluding Remarks

To answer the question what 'more information' in terms of information structures means several notions of informativeness have been developed. Blackwell's (1951) prominent sufficiency criterion is widely used but quite restrictive and subsequently weaker criteria have been used in economic applications.

---

[8] Actually, it is possible to take any a priori distribution for the counterexample to hold.

A more recent approach to compare information structures based on the sensitivity of the posterior state distributions with respect to changes in the signal has been followed by Eckwert and Zilcha (2004). Closely related is the work of Ganuza and Penalva (2006).

The relationship between the concepts of Blackwell, Eckwert and Zilcha as well as Ganuza and Penalva has been studied in this paper. For this purpose Blackwell's criterion was equivalently formulated using posterior distribution functions. It turned out that only in the case of binary information structures the Blackwell criterion implies the other ones which are in this case equivalent. In general there is no inclusion. Therefore relating results derived by the use of the two different approaches seems to be appropriate only in the context of the binary case.

The aim of further analysis should be to find a more suitable formalization of dispersion with respect to the observed signal to characterize more general Blackwell's notion of informativeness the way discussed in section 2.2 with the generation of state posteriors.

A Appendix

A.1 Binary Information Structures

In the case of binary information structures, the number of states of nature and signals is set to two. Such an information structure can be represented by the matrix

$$P = \begin{pmatrix} p_1 & 1-p_1 \\ 1-p_2 & p_2 \end{pmatrix} \tag{31}$$

with $0 \leq p_1, p_2 \leq 1$. Moreover, it suffices to restrict the parameter values to

$$p_1 + p_2 \geq 1, \tag{32}$$

because, if not the case, (32) is achievable by simply recoding the signal index. It should be noted that this assumption is equivalent to P satisfying the MLRP.

Lemma 3 *(Nermuth 1982) Let P and Q be two binary information structures satisfying* (32). *$P \succeq^{i} Q$ if and only if the following two conditions hold:*

$$\begin{aligned} (1-p_1)q_2 &\leq p_2(1-q_1) \\ (1-p_2)q_1 &\leq p_1(1-q_2) \end{aligned} \tag{33}$$

To get an insight to these conditions, suppose all entries in P and Q are non-zero, i.e. in both states there is for both signals at least a small probability to occur. Then (33) can be rearranged to

$$\frac{q_2}{1-q_1} \leq \frac{p_2}{1-p_1} \quad \text{and} \quad \frac{q_1}{1-q_2} \leq \frac{p_1}{1-p_2}. \tag{34}$$

Hence, under the more informative structure P the probability ratio of getting signal y_1 given state ω_1 and ω_2, respectively, is greater than under Q. The same holds for the probability ratio of getting signal y_2 given state ω_2 and ω_1, respectively. This means that the ratio of observing a signal when the corresponding and not corresponding state prevails is uniformly higher for the more informative structure.

A.2 The Stochastic Group

Poole (1995) points out that the set of non-singular stochastic $n \times n$ matrices together with standard matrix multiplication form a subgroup of the general linear group $GL(n)$. Note that in this context stochastic does not imply non-negative as associated with the term row-stochastic throughout the paper. Here it is just needed that all entries in a row sum to one. Let M be the set of all stochastic 2×2 matrices with trace greater than one, i.e. totally uninformative information structures, characterized by $p_1 + p_2 = 1$, i.e. identical conditional signal distributions, will not be considered.

Lemma 4 *The set M together with standard matrix multiplication is a subgroup of the stochastic group.*

Proof: Obviously, M is a subset of the set of non-singular stochastic 2×2 matrices and $I \in M$ for I denoting the identity matrix. For $P, Q \in M$ the product PQ is also an element of M, because

$$\begin{aligned} trace(PQ) &= 1-(p_1+p_2)+1-(q_1+q_2)+(p_1+p_2)(q_1+q_2) \\ &= 1+[(p_1+p_2)-1]\,[(q_1+q_2)-1] > 1, \end{aligned} \tag{35}$$

i.e. M is closed under multiplication. For any $P \in M$ one gets

$$trace(P^{-1}) = \frac{p_1 + p_2}{p_1 + p_2 - 1} > 1 \tag{36}$$

and therefore $P^{-1} \in M$. Thus M together with standard matrix multiplication is a subgroup of the stochastic group. □

References

Bickel, P. and E. Lehmann (1976). Descriptive statistics for nonparametric models. III. Dispersion. *Annals of Statistics 4*, 1139–1158.

Bielinska-Kwapisz, A. (2003). Sufficiency in Blackwell's theorem. *Mathematical Social Sciences 46*, 21–25.

Blackwell, D. (1951). The comparison of experiments. In J. Neymann (Ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 93–102. Berkeley: University of California Press.

Eckwert, B. and I. Zilcha (2004). Improvement in information: Income inequality and human capital formation. Manuscript.

Ganuza, J.-J. and J. Penalva (2006). On information and competition in private value auctions. Working Paper No. 937, Department of Economics and Business, Universitat Pompeu Fabra, Barcelona.

Hirshleifer, J. and J. G. Riley (1992). *The analytics of uncertainty and information.* Cambridge: Cambridge University Press.

Jewitt, I. (1997). Information and principal agent problems. Discussion Paper No. 97/414, University of Bristol, Bristol.

Kim, S. K. (1995). Efficiency of an information system in an agency model. *Econometrica 63*, 89–102.

Lehmann, E. (1988). Comparing location experiments. *Annals of Statistics 16*, 521–533.

Mac Lane, S. and G. Birkhoff (1967). *Algebra.* Toronto: Macmillan.

Marschak, J. and K. Miyasawa (1968). Economic comparability of information systems. *International Economic Review 9*, 137–174.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *Bell Journal of Economics 12*, 380–391.

Nermuth, M. (1982). *Information Structures in Economics.* Berlin: Springer-Verlag.

Nermuth, M. (1992). Different economic theories with the same formal strucutre: Risk, income inequality, information structures, etc. Working Paper No. 9207, Department of Economics, University of Vienna, Vienna.

Poole, D. G. (1995). The stochastic group. *The American Mathematical Monthly 102*, 708–801.