A two-person game of information transmission[☆]

Jerry R. Green[a,][*], Nancy L. Stokey[b]

[a]*Department of Economics, Harvard University, Littauer Center, Cambridge, MA 02138, USA*
[b]*Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 60637, USA*

Abstract

We consider a statistical decision problem faced by a two player organization whose members may not agree on outcome evaluations and prior probabilities. One player is specialized in gathering information and transmitting it to the other, who takes the decision. This process is modeled as a game. Qualitative properties of the equilibria are analyzed. The impact of improving the quality of available information on the equilibrium welfares of the two individuals is studied. Better information generally may not improve welfare. We give conditions under which it will.
© 2006 Published by Elsevier Inc.

1. Introduction

When a decision is made by a group rather than an individual, the twin tasks of acquiring information on the one hand, and choosing a course of action on the other, are often delegated to separate sets of individuals. If all members of the group share common evaluations of the outcomes and have identical prior beliefs, then there is no conflict between the information-gatherers and the action-takers. Information will be accurately transmitted by the former and optimally utilized by the latter.

Here we study the situation that arises when interests do not coincide. When interests diverge, complete transmission may result in actions that are suboptimal from the information-gatherers' point of view. The situation is one of partial conflict. We model it as a game in which each of the

---

two functions is executed by a single rational individual, neglecting conflicts among information-gatherers or among agents controlling different aspects of the group's action. We examine the Nash equilibria of the resulting two-person game. In particular, we look at the effect on the expected utilities of the two players of improvements in the available information.

The two individuals will be called the *agent* and the *principal*. Their joint decision problem is to choose an *action*, a_k, from the set $A = \{a_1, \ldots, a_K\}$. The von Neumann–Morgenstern utility levels of the two participants depend upon the chosen action and the realization of the *state of nature*, θ_m, from the set $\Theta = \{\theta_1, \ldots, \theta_M\}$. These utilities can be represented by $K \times M$ matrices $U = [u_{km}]$ and $U' = [u'_{km}]$ for the principal and the agent, respectively, where the elements are the utilities realized if a_k is chosen and θ_m occurs.

The agent receives an *observation* which is statistically related to the true state in Θ, and transmits the observation to the principal. He might not do so truthfully. There are N possible observations, y_n, in the set $Y = \{y_1, \ldots, y_N\}$. Allowing randomizations, his strategies can be represented by an $N \times N$ Markov matrix $R = [r_{nn'}]$, where $r_{nn'}$ is the probability that $y_{n'}$ is transmitted given that the actual observation is y_n.

The principal chooses the action $a_k \in A$ given that the observation $y_{n'}$ has been transmitted to him. Again, allowing randomization, his strategy is an $N \times K$ Markov matrix $Z = [z_{n'k}]$ where $z_{n'k}$ is the probability that a_k is chosen given that $y_{n'}$ was transmitted. [1]

The statistical relationship between states and observations is called the *information structure*. It is represented by an $M \times N$ Markov matrix $\Lambda = [\lambda_{mn}]$, where λ_{mn} is the probability that y_n is observed if the true state is θ_m. The interpretation of y_n depends on the *prior beliefs* of the individual in question. We allow different beliefs, $\pi = (\pi_1, \ldots, \pi_M) \in \Delta^M$ and $\pi' = (\pi'_1, \ldots, \pi'_M) \in \Delta^M$ for the principal and agent, respectively, where Δ^M is the set of all M-dimensional probability vectors. The principal's posterior probabilities, given an observation, can be derived from π and Λ by Bayes rule. These posteriors are denoted $(p_1^P, \ldots, p_N^P)$, where $p_n^P \in \Delta^M$ is his posterior if y_n is observed, for $n = 1, \ldots, N$. The probability of observing each y_n is also implied by π and Λ. Thus, we have a *distribution of the posterior* which is simply the measure over Δ^M assigning the corresponding weight to each of the p_n^P. A similar argument applies for the agent.

If the strategy choices are Z and R, the expected utilities for the principal and agent, respectively are

$$\text{tr } U \Pi \Lambda R Z \tag{1.1}$$

and

$$\text{tr } U' \Pi' \Lambda R Z, \tag{1.2}$$

where Π and Π' denotes the square matrices with the vectors π and π' on the diagonal and zeros elsewhere. [2]

---

[1] Whenever possible we will try to follow the convention of labeling typical actions and states with indices k and m, and the true and transmitted observations with n and n'.

[2] The principal's expected utility is

$$\text{EU}^P = \sum_m \pi_m \sum_n \lambda_{mn} \sum_{n'} r_{nn'} \sum_k z_{n'k} u_{km},$$

and a similar expression holds for the agent. The interpretation is straightforward. We simply sum up all the ways in which each action could occur given each possible state, by multiplying the conditional probabilities of observations given states, transmissions given observations, and actions given observations, and weighting by the prior probabilities.

In this paper we examine the Nash equilibria of this game. A pair of Markov matrices (Z, R) is a *Nash equilibrium* if Z maximizes (1.1) and R maximizes (1.2).

The main results of the paper can be viewed in the tradition of comparative statics. We are interested in the consequences of changes in the information structure (Y, Λ) on the equilibria of the game. Specifically, it is well-known that a partial ordering of information structures according to the criterion of *informativeness* can be given a precise mathematical characterization. This is an ordering based on single-person statistical decision theory. An information structure (Y, Λ) is said to be *more informative* than (Y', Λ') if, for any U and any π, the decision problem under the former has at least as high a value as that under the latter. Using the notation developed above, this can be restated as

$$\max_{Z} \operatorname{tr} U\Pi\Lambda Z \geqq \max_{Z'} \operatorname{tr} U\Pi\Lambda' Z',$$

where the maximum in each case is taken over all Markov matrices of the appropriate dimension. Blackwell [1] has shown the following:

> Λ is more informative than Λ' if and only if
> there exists a Markov matrix B such that $\Lambda' = \Lambda B$.

We want to study the relation between this condition and conditions sufficient for the improvement of the welfare of one or both of the players in our two-person organization. Because of the compounding of game theoretic aspects with the usual decision theoretic issues, the welfare of the two players may not be monotonic with respect to the quality of the information structure. Several types of complication arise.

First, as in most games, there may be multiple equilibria. We have found it hard to analyze all of them. However, a natural classification of equilibria can be given, and one type, which we call *partition equilibria,* have a rather regular behavior. Moreover, we will give some arguments to the effect that these equilibria have desirable properties, and are hence "more likely" to be observed.

Second, as in the case of general equilibrium theory, the set of equilibria is lower hemi-continuous with respect to changes in the parameters. Comparative static results therefore tend to be only local. Hence we focus on "small" improvements in the information structure, suitably defined.

Third, and finally, the comparative static results turn out to be different for the two players. For the agent, any small improvement in the information structure will improve his expected utility in a partition equilibrium. For the principal this may not be the case. His welfare can be guaranteed to be monotonic only when a very special kind of improvement in information is considered.

We define a *success-enhancing* improvement in information as one in which the probability that the observation is uninformative decreases, with a corresponding equiproportional increase in the probabilities of each of the other observations. If in the original information structure there is no such observation, that is if the posterior is unequal to the prior for every possible observation, then no success-enhancing improvements are possible. We show that small success-enhancing improvements in information necessarily improve the welfare of the principal at any partition equilibrium.

The remainder of the paper is organized as follows:

Section 2 covers the classification of types of equilibria and presents some genericity and stability-like arguments to bolster the case for considering partition equilibria. Section 3 contains the main comparative static results mentioned above. Section 4 contains several examples, primarily designed to illustrate directions in which our results cannot be extended.

2. Types of equilibria and their properties

2.1. Basic classification

We begin by examining some general features of the set of equilibria. First we need the following definition.

We will call an $M \times N'$ information structure Λ' a *partition* of Λ if $\Lambda' = \Lambda P D P'$, where P and P' are permutation matrices and D is an $N \times N'$ block diagonal Markov matrix in which each block has rank one. When Λ' has this form, it is as if there is a partition of the signal space Y. Under Λ', if signal value y_k occurs, the partition element containing y_k is reported. Thus P rearranges signal values so that those in a common partition element are clustered together; each block along the diagonal of D corresponds to the report for one partition element; and P' rearranges the new signal values in any arbitrary way.

Of prime importance in our later analysis are equilibria in which ΛR, the information the agent transmits, is a partition of Λ, the information he receives. These will be called partition equilibria. In addition, there are two types of non-partition equilibria, distinguished by whether the principal uses pure or mixed strategies. It is useful to begin with an example that illustrates all three types.

Example 1. There are two states, two actions, and two observations: $K = M = N = 2$, and

$$\Lambda = I, \quad U = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}, \quad U' = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}, \quad \pi = \pi' = (1/2, 1/2).$$

The agent knows the true state. Both parties prefer action a_2 in state θ_2. The principal strictly prefers action a_1 in state θ_1, but the agent is indifferent between the two actions.

One equilibrium of this game is the pair of strategies $Z = I$, $R = I$. Since the information transmitted by the agent is a partition (the complete refinement partition) of the space of observations, this is an example of a partition equilibrium.

Another equilibrium is one involving no transmission of information. This equilibrium is represented by any pair (Z, R) with

$$Z = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} \alpha & 1 - \alpha \\ \alpha & 1 - \alpha \end{pmatrix}, \quad \alpha \in [0, 1].$$

Note that for any α, $RZ = Z$. Clearly $\Lambda' = \Lambda R$ is a partition of Λ (the partition consisting of one set, equal to the whole space), so that this is another example of a partition equilibrium.

Another type of equilibrium is the pair

$$Z = I, \quad R = \begin{pmatrix} 1 - \varepsilon & \varepsilon \\ 0 & 1 \end{pmatrix}, \quad 0 < \varepsilon \leqslant 1/2.$$

Since $\Lambda' = \Lambda R$ is not a partition of Λ, this is not a partition equilibrium. Because the principal uses two distinct, nonrandomized actions, we call this a determinate action equilibrium. Note that randomization by the agent occurs only because in state θ_2 he is indifferent between the two actions a_1 and a_2. Clearly this situation is non-generic.

The last type of equilibrium is given by

$$Z = \begin{pmatrix} 1 & 0 \\ \delta & 1-\delta \end{pmatrix}, \quad R = \begin{pmatrix} 1/2 & 1/2 \\ 0 & 1 \end{pmatrix}, \quad \delta \in (0, 1).$$

As in the previous case, $\Lambda' = \Lambda R$ is not a partition of Λ. However, in contrast to the previous case, the principal is playing a mixed strategy. We call this a random action equilibrium. All equilibria of this form are non-robust to perturbations in U'. As long as Z has distinct rows the agent will not remain indifferent between actions, following a perturbation of U. In this sense, these equilibria are non-robust for the same reason as the previous type. However, in more general models the determinate action equilibria, with non-randomized actions by the principal in all cases, is always non-robust, whereas the random action equilibria may behave continuously in the parameters. These differences are explored more fully below.

Returning to the equilibria with both players randomizing, we note that they are all unstable in the sense that the agent has many optimal responses to Z, namely all R of the form

$$R = \begin{pmatrix} \alpha & 1-\alpha \\ 0 & 1 \end{pmatrix}, \quad \alpha \in [0, 1],$$

and the principal has many optimal responses to R, namely all Z of the form given above. But if the principal misperceives R even slightly, his optimal response is a non-randomized strategy, a Z matrix composed of zeros and ones, and the outcome would depart markedly from the equilibrium outcome.

2.2. Robustness of partition equilibria

In the rest of this section we will define partition, determinate action, and random action equilibria precisely, and argue that partition equilibria are robust in ways that the others are not. Specifically, we will show that determinate action equilibria are non-generic, and that random action equilibria are unstable against small perturbations in either player's strategy.

Formally, we will say that an equilibrium pair (Z, R) is a *partition equilibrium* if $\Lambda' = \Lambda R$ is a partition of Λ; a *determinate action equilibrium* if $\Lambda' = \Lambda R$ is not a partition of Λ, and each row of Z receiving positive weight under R has only a single positive element; a *random action equilibrium* if $\Lambda' = \Lambda R$ is not a partition of Λ, and some row of Z receiving positive weight under R has two or more nonzero entries.

Roughly speaking, we are presenting a "structural stability" argument to eliminate determinate action equilibria from consideration and a "dynamic stability" argument to eliminate random action equilibria. Of course, since we do not present any adjustment process, we do not actually have any dynamics. The objection to random action equilibria is only meant to be suggestive. Nevertheless, partition equilibria will generically pass both of these tests.

Theorem 2.1. *The set of all (U', Λ) for which there is any determinate action equilibrium is closed and null.*

Proof. The existence of a determinate action equilibrium requires that for some observation the agent is indifferent among some of the actions in A. It therefore suffices to show that the set of (U', Λ) for which this indifference holds is closed and null. But this property is obvious:

unless two rows of U' are identical (i.e. two actions are really the same) the set of posterior probabilities under which there is more than one optimal action is of lower dimensions than Δ^M. Hence the set of all Λ matrices for which these posteriors arise is null. Closedness is obvious. $\square$

Theorem 2.2. *Let* (Z, R) *be a random action equilibrium. Then generically in* (U, U', Λ) *there exists a sequence of Markov matrices* $\{R_v\}$ *converging to* R, *such that*

(i) *each* R_v *is an optimal response by the agent to* Z;
(ii) *the set of optimal responses by the principal to each* R_v *is bounded away from* Z.

Proof. Since (Z, R) is a random action equilibrium there exists $y_n \in Y$ and n_1, n_2 such that

(a) $r_{nn_1} > 0, r_{nn_2} > 0, n_1 \neq n_2$;
(b) the row vectors $z_{n_1}.$ and $z_{n_2}.$ are distinct; and
(c) the row vector $z_{n_1}.$ has at least two positive entries.

Since R is an optimal response to Z, it remains an optimal response if r_{nn_1} is increased by v and r_{nn_2} is decreased by v. Let R_v be the matrix with this change. Generically in Λ, this change in the agent's strategy alters the principal's posterior beliefs when the signal y_{n_1} is transmitted. Generically in U, this change in his posterior beliefs destroys the equality of the expected utilities under the actions represented in the mixture in the n_1^{st} row of Z. Hence Z is not a best response to R_v. Letting $v \to 0$ establishes the claim. $\square$

Theorem 2.3. *Generically in* (U, U', Λ), *if* (Z, R) *is a partition equilibrium, then,*

(i) *every row of* Z *for which the corresponding column of* R *has a positive entry is uniquely determined in the optimal response to* R;
(ii) R *is the unique optimal response to* Z *and is itself a partition.*

Proof. Obvious. $\square$

The main comparative static results of this paper apply to the generic instance of partition equilibria with the properties stated in Theorem 2.3. To delineate this class of equilibria more sharply, we give the following definition.

A partition equilibrium (Z, R) is called an *essential* equilibrium if the following two conditions hold:

(i) if Z' is an optimal response to R then $RZ' = RZ$;
(ii) if R' is an optimal response to Z then $R'Z = RZ$.

The idea of essential equilibria is that the strategies of each player are "essentially" unique, in the sense that choosing a different strategy from the optimal set does not alter the statistical relationship between the observations and the action taken. An essential equilibrium remains an equilibrium when either player chooses a different element in his set of optimal responses. Essential equilibria also possess a kind of "stability" in that they are robust to small deviations from optimal responses.

The distinction between partition equilibria in general and essential equilibria can be seen in the following example, to which we will return in Section 3.

Example 2. There are two actions, two states, and two signals, $K = M = N = 2$, and

$$U = U' = I,$$

$$\pi = (0.4, 0.6), \quad \pi' = (0.6, 0.4),$$

$$\Lambda = \begin{pmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{pmatrix}.$$

Consider $Z = R = I$. It is straightforward to verify that this is a partition equilibrium. The principal's posterior probabilities are

$$p_1^P = \left(\tfrac{1}{2}, \tfrac{1}{2}\right), \quad p_2^P = \left(\tfrac{4}{13}, \tfrac{9}{13}\right),$$

and, symmetrically, the agent's are

$$p_1^A = \left(\tfrac{9}{13}, \tfrac{4}{13}\right), \quad p_2^A = \left(\tfrac{1}{2}, \tfrac{1}{2}\right).$$

Therefore, the agent is indifferent between both actions when he receives the observation y_2, and the principal is indifferent when the agent transmits y_1. The choices of $Z = R = I$ are mutually fortuitous, and neither requirement in the definition of an essential equilibrium holds. The non-genericity of non-essential partition equilibria is responsible for their peculiar comparative static properties, as we will see below.

3. Improvements in the information structure

In this section we present the main comparative static results of this paper. We ask the question: When can one be sure, independent of a knowledge of the preferences and beliefs of the two individuals, that one information structure is better than another in the sense of providing a higher level of expected utility in equilibrium? The answer depends on whose welfare is being considered. Broadly speaking we find that the agent benefits from any improvement in the information structure. The principal, however, can well be hurt. Only for one very special, though interesting, type of improvement can we be sure that the principal benefits.

One further qualification is important to emphasize. As in many games, and as we have seen in the examples of Section 2, there are often multiple equilibria. Because they can be regarded as fixed points of a suitable mapping, they are continuous in the parameters of the problems for almost all parameter values. However, at some critical points the set of equilibria changes radically. Non-essential partition equilibria, for example, are likely to occur at such points.

For this reason our results are "local" in nature. We define a concept of "small" changes in the information structure. The comparative static results described above apply to changes in information that are sufficiently small, at an equilibrium that moves continuously in this change.

The concept of small changes or, mathematically, a topology on the space of information structures, is given by the following definition of convergence. It is natural to say that a sequence of information structures $\{\Lambda_v\}$ *converges* to Λ_0 if for any decision maker with utility matrix U and prior π, the sequence of values $\{V(U, \pi, \Lambda_v) = \max_{D \in M} \operatorname{tr} U \Pi \Lambda_v D\}$ converges to $V(U, \pi, \Lambda_0)$. Obviously this is equivalent to the weak convergence of the distribution of the posteriors for any strictly positive prior. Information structures representing a *small improvement* in information

from Λ_0 are those in a neighborhood of Λ_0 which are also more informative in the sense of Blackwell.

It is important to point out that the dimensionality of the likelihood matrices, that is to say the number of possible observations N_v, is not held constant. We are able to compare information structures in which the qualitative nature of the signals are quite different.

We will now show that essential equilibria have an invariance property that is responsible for the comparative static results that we will obtain.

Theorem 3.1. *Let $\{\Lambda_v\}$ be a sequence of information structures converging to Λ_0, and let (Z_0, R_0) be an essential equilibrium for Λ_0. There exists a sequence $\{(Z_v, R_v)\}$ of equilibria corresponding to $\{\Lambda_v\}$ with the property that $\{\Lambda_v R_v\}$ converges to $\Lambda_0 R_0$.*

Proof. The proof is by construction of a sequence $\{(Z_v, R_v)\}$ having the required properties.

For each $a_k \in A$, let $B_k \subseteq \Delta^M$ be the set of posterior beliefs for which the principal strictly prefers a_k to the other actions in A. Note that these sets are disjoint. Since (Z_0, R_0) is essential, the principal's posterior given any signal from $\Lambda_0 R_0$ lies in the interior of one of the sets B_k.

Let $A_0 \subseteq A$ be the subset of actions receiving positive weight under Z_0, and for each $a_k \in A_0$, let $C_k \subseteq \Delta^M$ be the set of posterior beliefs for which the agent strictly prefers a_k to the other actions in A_0. Since (Z_0, R_0) is essential, the agent's posterior given any signal from Λ_0 lies in the interior of one of the sets C_k.

Let $\hat{Z}$ be any matrix of appropriate dimension whose distinct rows are precisely the distinct rows of Z_0 receiving positive weight under R_0. Select a subset J of these rows containing all the distinct rows of $\hat{Z}$ and no duplicates. Let R_v be any response to $\hat{Z}$ that is optimal for the agent, given the information structure Λ_v and subject to the constraint that only signals in J are transmitted. By construction R_v is also an unconstrained optimal response for the agent, given Λ_v and $\hat{Z}$.

Since $\{\Lambda_v\}$ converges to Λ_0, as $v \to \infty$, with probability approaching one the agent's posteriors given the signals in Λ_v lie in the interiors of the same sets C_k as they do under Λ_0. Hence $\{\Lambda_v R_v\}$ converges to $\Lambda_0 R_0$.

Finally, note that by construction the number of distinct signals transmitted under $\Lambda_v R_v$ is the same as the number transmitted under $\Lambda_0 R_0$. Hence for v sufficiently large the principal's posterior under any signal from $\Lambda_v R_v$ lies the interior of the set B_k corresponding to the action selected by $\Lambda_v R_v \hat{Z}$. Hence $Z_v = \hat{Z}$ is an optimal response for the principal, and the sequence $\{(Z_v, R_v)\}$ has the required properties. $\quad\square$

Theorem 3.2. *Let (Z_0, R_0) be an essential equilibrium for the information structure Λ_0. Let $\{\Lambda_v\}$ be a sequence of information structures, each element of which is more informative than Λ_0 (in the sense of Blackwell), and such that $\{\Lambda_v\}$ converges to Λ_0. Let $\{(Z_v, R_v)\}$ be the sequence of equilibria whose existence is established in Theorem 3.1. Then, for v sufficiently large, the agent is better off under Λ_v with the equilibrium (Z_v, R_v) than under Λ_v with the equilibrium (Z_0, R_0).*

Proof. As the rows of Z_v are precisely the distinct rows of Z_0, by construction the agent is facing a fixed decision problem and hence his welfare cannot diminish under any improvement in the sense of Blackwell. $\quad\square$

The next example illustrates the role of the hypothesis that the equilibrium is essential.

Example 3. There are four actions, four states, and three or four signals, $K = M = 4$, $N_0 = 3$, and $N_\varepsilon = 4$. The utilities are

$$
U = \begin{pmatrix} 1 & -b & b & 0 \\ 0 & a & -a & 0 \\ 0 & b & -b & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad U' = \begin{pmatrix} 0 & d & -d & 0 \\ 0 & c & -c & 0 \\ 0 & -d & d & 0 \\ -e & -e & -e & -e \end{pmatrix},
$$

where

$$
0 < a < b, \quad 0 < c < d, \quad 0 < e.
$$

Both agents have a uniform prior, $\pi = \pi' = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \right)$, and the information structures are

$$
\Lambda_\varepsilon = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & (1+\varepsilon)/2 & (1-\varepsilon)/2 & 0 \\ 0 & (1-\varepsilon)/2 & (1+\varepsilon)/2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \Lambda_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
$$

States θ_1 and θ_4 are perfectly revealed under both information structures. The more informative system Λ_ε is useful only in distinguishing state θ_2 from θ_3.

Under Λ_0 the posteriors are

$$
p_{10} = (1, 0, 0, 0), \quad p_{20} = (0, 1/2, 1/2, 0), \quad p_{30} = (0, 0, 0, 1),
$$

so both the principal and the agent are indifferent among the actions $\{a_1, a_2, a_3\}$ if y_2 is observed: each is a fair bet. Thus,

$$
\hat{R} = I, \quad \hat{Z} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},
$$

is a partition equilibrium. The agent accurately reports the observation, and the principal chooses a_n if y_n is reported, for $n = 1, 2, 3$. The principal's expected utility is $\frac{1}{2}$ and the agent's is 0.

This equilibrium is not essential, however. The principal is indifferent between $\hat{Z}$ and the strategy of taking action a_4 all the time,

$$
Z_0 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
$$

There is another partition equilibrium under Λ_0, the no communication equilibrium. The principal uses the strategy Z_0 and the agent uses any Markov matrix R_0 with rank one. The principal's expected utility is again $\frac{1}{2}$, but the agent's is $-e$. This equilibrium is essential: $RZ_0 = R_0 Z_0$ for any Markov matrix R, so the agent cannot affect the outcome; and the principal's best response to R_0 is unique.

For $\varepsilon > 0$, the only equilibrium is the no communication equilibrium. The principal uses the strategy Z_ε that puts probability one on a_4 in response to any report, and the agent uses

any Markov matrix R_ε with rank one. The principal's expected utility is again $\frac{1}{2}$ and the agent's is $-e$.

To see why there are no other equilibria note that the posteriors under Λ_ε are

$$p_{1\varepsilon} = (1, 0, 0, 0), \quad p_{2\varepsilon} = (0, (1 + \varepsilon)/2, (1 - \varepsilon)/2, 0),$$

$$p_{4\varepsilon} = (0, 0, 0, 1), \quad p_{3\varepsilon} = (0, (1 - \varepsilon)/2, (1 + \varepsilon)/2, 0).$$

Thus, the agent strictly prefers a_1 to a_2 if y_2 is observed, and strictly prefers a_3 to a_2 if y_3 is observed: the former involves increasing the size of a favorable gamble, and the latter involves comparing an unfavorable gamble with one that is favorable. Suppose the principal were to use the strategy of playing a_1 if y_1 is reported, a_2 if y_2 or y_3 is reported, and a_3 if y_4 is reported, which under Λ_ε and accurate reporting is the analog of $\hat{Z}$. The agent would never report y_2 or y_3. Instead he would report y_1 or y_4, which gives him an expected utility of $d\varepsilon/2$ instead of the 0 he gets by reporting accurately. This would reduce the principal's expected utility, however, and since the principal can always guarantee himself an expected utility of $\frac{1}{2}$ by ignoring the reports and taking the action a_4, he would do so. Therefore, the analog of the $\left(\hat{R}, \hat{Z}\right)$ equilibrium does not exist for $\varepsilon > 0$, and a sequence of the type described in Theorem 3.1 does not exist.

We now consider the principal's welfare. Further conditions are required to guarantee that better information raises the principal's equilibrium expected utility. A modification of Example 2 is useful in gaining insights to the results.

Example 4. As in Example 2, $K = M = 2$, and for Λ_0 the number of distinct signals is $N_0 = 2$. However, for all υ there are three possible observations, with the likelihood matrices

$$\Lambda_\upsilon = \begin{pmatrix} .6 - \varepsilon_\upsilon & 2\varepsilon_\upsilon & .4 - \varepsilon_\upsilon \\ .4 - \varepsilon_\upsilon & 2\varepsilon_\upsilon & .6 - \varepsilon_\upsilon \end{pmatrix},$$

where $\{\varepsilon_\upsilon\}$ converges to zero. Clearly $\{\Lambda_\upsilon\}$ converges to the information structure of Example 2,

$$\Lambda_0 = \begin{pmatrix} .6 & .4 \\ .4 & .6 \end{pmatrix}.$$

Moreover $\{\Lambda_\upsilon\}$ is a sequence of improvements, in the sense of Blackwell. We modify the priors so they are slightly more diffuse

$$\pi = (0.45, 0.55), \quad \pi' = (0.55, 0.45).$$

The other data of the example are unchanged, $U = U' = I$. Let y_1 and y_2 denote the signals under Λ_0, and $\bar{y}_1, \bar{y}_2, \bar{y}_3$ denote those under Λ_υ.

Under Λ_0 it is straightforward to verify that $Z = R = I$ is an essential partition equilibrium. It results in an expected utility of 0.6 to each player.

For ε_v sufficiently small, the game with information structure Λ_v has an essentially unique partition equilibrium (other than the no-transmission equilibrium). The agent transmits the partition $\{\bar{y}_1, \bar{y}_2\}, \{\bar{y}_3\}$, and in response the principal chooses a_1 and a_2, respectively. Thus action a_1 is taken when y_2 is observed, as the agent prefers. The expected utilities at this equilibrium are $0.6 - \varepsilon_v$ and $0.6 + \varepsilon_v$ for the principal and the agent, respectively. Although the principal would

rather have the agent transmit the partition $\{\overline{y}_1\}$, $\{\overline{y}_2, \overline{y}_3\}$, there is nothing he can do to enforce this. Complete communication is not an equilibrium because the principal would choose a_2 after y_2, making complete communication irrational for the agent. In this example the better information structure entails a positive probability of a signal that causes the two players to disagree. The principal loses because at equilibrium the information is used by the agent in a way opposite to what the principal would like.

Our positive comparative static results rely on a condition that we will call *success-enhancing*.

The motivation for examining success enhancing improvements is that there are many situations where one hopes to receive an informative observation but in fact nothing happens. Either the experiment "fails" or the outcome is not available soon enough to be useful in making the decision. Since many improvements in information reduce the failure rate or cut the average delay time without affecting the quality of the experimental procedure itself, these results are of considerable interest. The formal definition is as follows.

We will say that Λ is in *standard form* if

$$
\Lambda = \begin{pmatrix} \alpha \\ \vdots & (1-\alpha)\,\Gamma \\ \alpha \end{pmatrix},
$$

where $\alpha \geqslant 0$, Γ is an $M \times (N-1)$ Markov matrix, and no two columns of Γ are proportional to each other.

The first signal of an information structure in standard form represents the totally uninformative observation: "dropping the test tube." A success-enhancing improvement lowers the probability of this observation and raises all others proportionally.

We will say that Λ is a *success-enhancing improvement* of Λ' if

$$
\Lambda = \begin{pmatrix} \alpha \\ \vdots & (1-\alpha)\,\Gamma \\ \alpha \end{pmatrix}, \quad \Lambda' = \begin{pmatrix} \alpha' \\ \vdots & (1-\alpha')\,\Gamma P \\ \alpha' \end{pmatrix},
$$

where

$$
0 \leqslant \alpha < \alpha' \leqslant 1 \quad \text{and } P \text{ is a permutation matrix.}
$$

Note that Λ' and Λ' are related by the garbling matrix

$$
B = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \beta & & & \\ \vdots & & (1-\beta)\,P & \\ \beta & & & \end{pmatrix} \quad \text{where } \beta \equiv \frac{\alpha' - \alpha}{1 - \alpha}.
$$

Finally, a *small success-enhancing improvement* is one in which α' and α are close.

Whenever Λ is a success-enhancing improvement of Λ' we can reorder the columns of Λ' so that the permutation matrix P referred to in the above definition is the identity. For the rest of this section we will suppose that this is the case, as this in no way changes the structure of the game.

With this normalization we can write

$$\Lambda' = \Lambda D,$$

where D is the diagonal matrix whose diagonal entries are

$$d_1 = \frac{\alpha'}{\alpha}, \quad d_n = \frac{1 - \alpha'}{1 - \alpha} \quad \text{for } n = 2, \ldots, N. \tag{3.1}$$

We will make use of this representation in the proof of the main theorem, which follows.

Theorem 3.3. *Let (Z, R) be an essential equilibrium for Λ' and let Λ be a small success-enhancing improvement of Λ'. Then (Z, R) remains an equilibrium for Λ and the principal's expected utility cannot decrease.*

Proof. That (Z, R) remains an equilibrium follows from an argument parallel to that used in the proof of Theorem 3.1. The assertion that the principal's expected utility cannot decrease will be proven using the special structure of success-enhancing improvements. We will express Λ' and Λ in standard form and note that $\Lambda' = \Lambda D$, where the elements of D are given by (3.1). The principal's expected utility under Λ is $\operatorname{tr} U\Pi\Lambda RZ$ and under Λ' it is $\operatorname{tr} U\Pi\Lambda DRZ$. Thus, the gain in going from Λ' to Λ is

$$\Delta = \operatorname{tr} RZU\Pi\Lambda (I - D). \tag{3.2}$$

This quantity will be proven to be necessarily non-negative.

Using (3.1) we see that

$$I - D = (1 - d_n) I - (d_1 - d_n) C$$

$$= \frac{\alpha' - \alpha}{1 - \alpha} I - \frac{\alpha' - \alpha}{\alpha (1 - \alpha)} C,$$

where

$$C \equiv \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & 0 & \\ 0 & & & 0 \end{pmatrix}.$$

Substituting into (3.2) we have

$$\Delta = \frac{\alpha' - \alpha}{1 - \alpha} \left[\operatorname{tr} RZU\Pi\Lambda - \frac{1}{\alpha} \operatorname{tr} RZU\Pi\Lambda C \right]. \tag{3.3}$$

Since Λ is in standard form its first column is the constant α, and we have

$$\Lambda C = \alpha J, \tag{3.4}$$

where J is the $M \times N$ Markov matrix

$$J \equiv \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}.$$

Substituting (3.4) into (3.3) we obtain

$$\Delta = \frac{\alpha' - \alpha}{1 - \alpha} \left[\operatorname{tr} RZU\Pi\Lambda - \operatorname{tr} RZU\Pi J \right]$$

$$= \frac{\alpha' - \alpha}{1 - \alpha} \left[\operatorname{tr} U\Pi\Lambda RZ - \operatorname{tr} U\Pi J RZ \right], \tag{3.5}$$

where the second line uses the commutativity of matrix multiplication under the trace. The two terms in brackets in (3.5) have straightforward interpretations. The first is the equilibrium expected value of the principal under the better information structure. The second is the expected utility he would obtain if he used the action matrix Z in a decision problem with information structure JR. But this information structure is totally uninformative: it is an $M \times N$ Markov matrix that contains the first row of R repeated M times. Therefore, the second term in brackets is at most the expected value the principal could obtain by optimizing in a situation where no information is available. This cannot exceed the first term, so the expression in brackets is nonnegative. Since $\alpha' - \alpha \geqslant 0$, it follows that $\Delta \geqslant 0$.

We note that Δ is zero only if the principal is indifferent, in equilibrium, between Z and a rank one matrix composed of a repeated row. Otherwise Δ is strictly positive. $\square$

4. Comments and further examples

In this section we gather a few comments showing why the results above cannot be strengthened and addressing some conjectures about the qualitative nature of the Nash equilibria.

1. Success-enhancing improvements in information have the property that the set of posterior beliefs that can arise after seeing the observation remains fixed. One might imagine that this property alone is responsible for the beneficial nature of the change.

An improvement in information from Λ' to Λ can be called *posterior-preserving* if $\Lambda' = \Lambda D$, where D is a diagonal matrix, and $\Lambda' = \Lambda B$, where B is a Markov matrix. Note that a non-informative signal may not exist. If (Z, R) is an equilibrium for the information structure Λ', then we know that R is among the agent's best responses to Z under Λ. The following example shows, however, that the principal's welfare may decrease if he plays Z and, moreover, that there may be no possibility for him to achieve the former level of utility.

Let $K = 2$, $M = 3$, and $N = 4$. Consider the information structures

$$\Lambda = \begin{pmatrix} .8 & 0 & 0 & .2 \\ 0 & .9 & 0 & .1 \\ 0 & 0 & .8 & .2 \end{pmatrix}, \quad \Lambda' = \begin{pmatrix} .6 & 0 & 0 & .4 \\ 0 & .8 & 0 & .2 \\ 0 & 0 & .6 & .4 \end{pmatrix}.$$

It is easy to see that Λ is a posterior-preserving improvement of Λ'. The first three signals in either case are perfect predictors of the state, while the fourth carries some information but does not limit the set of states that are possible.

Let the utilities and priors be such that

$$U\Pi = \begin{pmatrix} -1 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \quad U'\Pi' = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}.$$

The finest partition equilibria convey $\{y_1, y_2\}, \{y_3, y_4\}$. For example, the strategies

$$
Z = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}
$$

constitute an equilibrium. In it expected utility for the principal is .2 under Λ' but .1 under Λ. More information has therefore been harmful.

Notice that even though Λ is strictly more informative that Λ', ΛR is non-comparable to $\Lambda' R$. For this particular utility function it is worse.

2. Even if the improvement in information is success enhancing, a discrete change from Λ' to Λ may be such that (Z, R) is no longer an equilibrium. R remains a best response to Z, but Z may not be best against R.

An example, available from the authors on request, shows that *all* the equilibria under Λ may be inferior for the principal to a given partition equilibrium for Λ'. Theorem 3.3 relies on the changes being small enough so that (Z, R) persists as an equilibrium.

3. We investigated the conjecture that the common refinement of the partitions implicit in two partition equilibria always corresponds to another partition equilibrium. A counterexample to this conjecture is also available on request.

4. Crawford and Sobel [3] show that some more general results can be attained if a restricted set of decision problems is considered. When utilities are concave in actions and states and both the state space and the signal space are one-dimensional, all equilibria are partition equilibria. They use this property to derive the comparative statics of both principal and agent welfare with respect to changes in their preferences, obtaining specific results for changes that make their preferences more similar to each other in a sense that they define.

5. Finally, it should be emphasized that the main results of this paper are crucially dependent on the finiteness of the set of possible actions. These results are local in nature, as noted above in point (2). The structure of our model is such that within a neighborhood of a given information structure, partition equilibria are locally constant with respect to success-enhancing or posterior preserving improvements. This enables us to evaluate welfare changes by examining the effect of the improved information in a fixed equilibrium.

If there were a continuum of actions the neighborhood of local constancy might vanish. Changing information would induce locally continuous shifts in the equilibria. Welfare effects would then depend upon the nature of these shifts, as well as on the difference in the quality of the information.

Postlewaite [7] has provided us with an example of a game with a unique partition equilibrium in which the principal's welfare declines, for this reason, in response to a success-enhancing improvement in the agent's information.

5. Conclusions

We have examined a simple two-person game designed to represent the separation of functions in an organization. It has been argued that although this game may have multiple equilibria, there is one type of particular note. In analyzing the comparative statics of individual welfare with respect to improvements in information, we have concentrated on this type of equilibrium.

In general, improvements in information may be harmful for one or even both players. We therefore tried to find restrictions on the nature of the improvement in information that imply that it is surely beneficial.

For large shifts in the information structure, nothing can be said, in general. Locally, an arbitrary improvement in the information structure will generically benefit the agent, but the principal may be hurt. To guarantee that neither player is hurt by a small improvement in information, the change must be "success-enhancing." That is, it must represent a decrease in the probability of receiving an uninformative observation and, correspondingly, proportional increases in the probabilities of receiving all other observations.

There are many possible extension of this model. We will mention only two of them here.

Our analysis concentrated on restricting the information structure. An alternative is to look for restrictions on utilities and priors. In this regard the paper by Crawford and Sobel [3] cited above is relevant.

Our model is related to, but distinct from, the principal-agent problem that has been widely discussed in the literature. There the agent plays the role of both information gatherer and decision maker. The principal is present only to help offset risks by making contingent payments of a transferable resource. We have no such resource, the essential feature of our model resting in the separation of the information-gathering and decision-making functions within the organization. The possibility of making such conditional payments would add an entirely new dimension to the analysis. The principal might, for example, set up a payment schedule that would coax a more accurate transmission out of the agent. Paralleling the principal-agent literature, it is probably best to model this as the Stackelberg, rather than Nash, equilibrium of a game in which the principal is the leader. We begin the study of this solution concept in a companion paper, Green and Stokey [5], retaining the structure presented here in all other respects—including the absence of transferable utility.

Acknowledgments

References

[1] D. Blackwell, Comparison of experiments, in: J. Neymann (Ed.), Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, CA, 1951, pp. 93–102.

[2] H. Bohnenblust, L. Shapley, S. Sherman, Reconnaissance in Game Theory, RM-208, Rand Corporation, Santa Monica, CA, 1949.

[3] V. Crawford, J. Sobel, Strategic information transmission, Econometrica 50 (1982) 1431–1451.

[4] J. Farrell, M. Rabin, Cheap talk, J. Econ. Perspect. 10 (1996) 103–118.

[5] J. Green, N. Stokey, Value of Information in the Delegation Problem, revised October 1981, Mimeo. (First version: HIER discussion paper no. 776, July 1980.)

[6] J. Marschak, K. Miasawa, Economic comparability of information systems, Int. Econ. Rev. 9 (1968) 137–174.

[7] A. Postlewaite, Private communication to the authors, 1980.