HAGEN LINDSTÄDT

MORE NONCONCAVITIES IN INFORMATION PROCESSING FUNCTIONS

ABSTRACT. The productivity of (human) information processing as an economic activity is a question that is raising some interest. Using Marschak's evaluation framework, Radner and Stiglitz have shown that, under certain conditions, the production function of this activity has increasing marginal returns in its initial stage. This paper shows that, under slightly different conditions, this information processing function has repeated convexities with ongoing processing activity. Even for smooth changes in the signals' likelihoods, the function is only piecewise smooth with non-differentiable convexities at points of conditional changes of action. For linear likelihood functions the processing value proves to be piecewise linear with convexities at these levels.

KEY WORDS: Value of Information, Information Structures, Information Processing Functions, Information Economics, Decision

1. INTRODUCTION

In today's economics, the fact that information is a major factor of production and of increasing importance as a source of competitive advantage is a generally accepted truism. It is only possible to evaluate information economically when a specific usage for the information is assumed. This usage typically applies to a certain decision problem, in which the information might change the decision makers' subjective prior probabilities for given environmental states. Furthermore, the evaluation has to be possible from a prior perspective – that is, before actually knowing the signal – in order to be useful when deciding about acquiring and processing the information. The standard framework for this evaluation is briefly stated in Section 2.

The analysis of the productivity of information as a factor of production and of information processing as an economic activity leads to a number of remarkable results. Just to recall a few, investments in information are often irreversible, which sometimes makes trading information difficult (Arrow 1974). Additionally, information

Theory and Decision **51:** 351–365, 2001.

does not disappear when used, so it can be used more than once, although its value might change while using it. Furthermore, information can be used for different tasks at the same time. These features demonstrate the fact that there are increasing marginal returns to information as a result of its multiple usage.

Besides these considerations, Radner and Stiglitz (1984) have shown that under certain conditions there are increasing marginal returns to a small amount of information even if it is only used for one specific decision problem. As will be seen later, initial information processing is not the only case for convexities when using information only once, i.e., in a single decision problem. For slightly different conditions than those of Radner and Stiglitz, the activity of information acquisition and processing has repeated convexities even when continued.

2. DISCUSSION OF THE RELEVANT LITERATURE

2.1. *Marschak's value of information and Blackwell's condition for the informativeness of information structures*

The standard in economic valuation of information is Marschak's model (Marschak 1954, 1959). It evaluates information with the gain a decision maker can extract from it in a specific decision problem due to a change in his subjective probabilities for given environmental states. For simplicity, we restrict ourselves to a risk-neutral decision maker, a finite number of environmental states S_s and alternative courses of actions A_a to choose from. When deciding about acquiring costly information like a market report, the decision maker has to value the information on a prior basis, i.e., without knowing its content (what the market report actually says). This is achieved by distinguishing between the information I (the report) and the signals I_i, representing alternative versions of its content which the decision maker believes to be possible (e.g., growing, stagnating or shrinking revenue for a very simple market report).

The prior value of the information $I = (I_1, \dots, I_I)$ before knowing the actual signal I_i is the expected additional payoff the decision maker can get from the decision due to possessing the information, i.e., knowing the signal. To calculate this payoff, he has to judge the likelihoods $\mathrm{p}(I_i|S_s)$ for the signals conditional on the possible states and work out the conditional probabilities $p(S_s|I_i)$ using Bayes' well-known equation. The resulting value of information equals the difference in expected payoff for the decision maker with and without the information (see (1)). The Markov matrix $(p(I_i|S_s))_{is}$ of likelihoods is called information structure. If U_{as} is the payoff of action A_a in state S_s in decision problem D and a^* indexes the action with maximal expected payoff without the information, then the value of information is usually defined as:

$$
\begin{aligned}
\mathrm{IV}_D((\mathrm{p}(I_i|S_s))_{is}) :=& \sum_{i=1}^{I} \mathrm{p}(I_i) \times \left[\max_{a=1,\dots,A} \left\{ \sum_{s=1}^{S} \mathrm{p}(S_s|I_i) \times U_{as} \right\} \right. \\
& \left. - \sum_{s=1}^{S} \mathrm{p}(S_s|I_i) \times U_{a^*s} \right] \\
=& \left[\sum_{i=1}^{I} \mathrm{p}(I_i) \times \max_{a=1,\dots,A} \left\{ \sum_{s=1}^{S} \mathrm{p}(S_s|I_i) \times U_{as} \right\} \right] \\
& - \left[\sum_{i=1}^{I} \left(\sum_{s=1}^{S} \mathrm{p}(I_i) \times \mathrm{p}(S_s|I_i) \times U_{a^*s} \right) \right] \\
=& \left[\sum_{i=1}^{I} \mathrm{p}(I_i) \times \max_{a=1,\dots,A} \left\{ \sum_{s=1}^{S} \mathrm{p}(S_s|I_i) \times U_{as} \right\} \right] \\
& - \left[\sum_{s=1}^{S} \left(\sum_{i=1}^{I} \mathrm{p}(S_s|I_i) \times \mathrm{p}(I_i) \right) \times U_{a^*s} \right] \\
=& \left[\sum_{i=1}^{I} \mathrm{p}(I_i) \times \max_{a=1,\dots,A} \left\{ \sum_{s=1}^{S} \mathrm{p}(S_s|I_i) \times U_{as} \right\} \right] \\
& - \left[\sum_{s=1}^{S} \mathrm{p}(S_s) \times U_{a^*s} \right] \qquad (1)
\end{aligned}
$$

When evaluating information processing, the natural question arises as to how to compare two given information structures (matrices P and Q of likelihoods) with respect to their informativeness. In general, we know there is no way to measure the quantity of information as a real number, and, of two given information structures, neither one might be more informative. Still, it is a well-known result attributable to Blackwell (1951), that for any two given information structures P and Q the statement '*P is more informative than Q*' is equivalent to the existence of a matrix M such that

$$P \times M = Q \tag{2}$$

'More informative than' means that for any utility function and decision problem the value of information in P is at least equal to the value of information in Q.[1]

It seems natural to restrict the term *information processing* to activities that lead to information structures becoming increasingly more informative. In other words, a matrix valued *information structure function*

$$\mathrm{IF}_P : t \to (\mathrm{p}_t(I_i|S_s))_{is} =: P_t \tag{3}$$

that maps real-valued activity levels on information structures is only said to model information processing when for $t \geq r$, P_t is more informative than P_r, that is, when for each r and t with $r \leq t$, there exists an M_{rt} with

$$(\mathrm{p}_t(I_i|S_s))_{is} \times M_{rt} = (\mathrm{p}_r(I_i|S_s))_{is} \tag{4}$$

2.2. *The Radner-Stiglitz result: The nonconcavity of initial information processing*

When interpreting processing of information as an economic activity that may add value in a decision problem, a natural question about the input/output-relationship of this activity arises, i.e., about its production function. Here, the independent variable is the activity level of processing information, while the dependent variable is the value of the information processed. The decision problem D, the decision maker's preferences (here: risk-neutral payoff maximizer), and a specific strategy for processing information have to be clear in order to make sense. By information processing strategy, we mean a plan or procedure about which a 'piece of information' is processed at a certain activity level. This production function with a real-valued activity level t will be called *information processing function*.[2]

Radner and Stiglitz (1984) have shown a substantial and very interesting result for information processing functions with real-valued arguments (activity levels) when processing information is costly: under certain conditions that will be discussed below, the net marginal value of costly information processing is negative near the origin of zero activity, i.e., in a neighborhood of the origin. Radner and Stiglitz conclude that there have to be increasing marginal returns to information processing over some range of the parameter as soon as there is some amount of information processing with a positive net value.

Arrow (1985) has put this result succintly: given the Radner–Stiglitz conditions, '*a little information is never worth the cost*'. Since Arrow (1985) drew attention to the fact that the result strongly depends on its assumptions, by giving examples that neither fulfill the conditions nor the result, it seems worthwhile to take a closer look at the four main assumptions from which Radner and Stiglitz were able to derive their strong result.[3]

First, Radner and Stiglitz assume an infinite number of actions that can be indexed with a multidimensional, real-valued parameter. An example of a decision problem with such an action set is the mixture of a securities portfolio. For k securities, the courses of action are the possible mixtures, a set that can be conveniently indexed with numbers from $[0;1]^k$ that add up to one — a $(k-1)$-dimensional subset of $[0;1]^k$.

They next assume payoff to be a continuous function in the parameterization of the actions, thus introducing a neighboring structure among courses of action that, apart from the parameters, also applies to the payoff. Once a decision problem has an infinite number of alternative actions indexed by real-valued parameters, it is natural to assume continuity: if any functional dependency between payoff and parameterization of alternative actions were allowed, one could hardly make statements about payoff shifts with changing probabilities and, consequently, about the value of information, since slight changes in the optimal solution could easily imply a completely different payoff.

The third assumption is made on the informativeness of the information structure $(\mathrm{p}_t(I_i|S_s))_{is}$ at $t = 0$, i.e., before actually starting to process information: at t=0, the information structure is assumed to be *completely non-informative* for all signals I_i, non-informative meaning that the likelihoods $\mathrm{p}_0(I_i|S_s)$ at $t = 0$ are independent from s for all signals I_i. As indicated by the term, completely non-informative information structures bear no information that could change the states' probabilities $\mathrm{p}(S_s)$ when knowing the signal. Especially important, the value of non-informative information structures (before costs) is always zero. Assuming non-informativeness means information processing actually starts without information that is not already contained in the probabilities $\mathrm{p}(S_s)$.

Accordingly, the fourth assumption rules out information processing procedures with sudden jumps or discontinuities in informativeness at the critical start: Radner and Stiglitz assume the information structure function IF_P: $t \rightarrow (\mathrm{p}_t(I_i|S_s))_{is}$ to be differentiable in $t = 0$. Consequently, there are no 'sudden insights' or jumps in informativeness. Instead, information processing is assumed to happen smoothly. Both the third and fourth assumption seem to be a natural condition for deriving the convexity result, since for information processing procedures deviating from these assumptions, general results seem hard to obtain. This does not, however, mean that real-world information processing might not proceed with sudden, non-differentiable insights in many cases.

3. A MODEL FOR THE VALUATION OF INFORMATION PROCESSING

To model the production function of information processing, we map a real-valued parameter representing the information processing activity to its value, i.e., to the value of the information structure that is connected with that activity. As a restriction, we only allow information structures that are successively more informative, i.e., we restrict processing to information structure functions $\mathrm{IF}_P : t \rightarrow (\mathrm{p}_t(I_i|S_s))_{is}$ as in Equation (3) that fulfill Equation (4). Additionally, IF_P is assumed to be continuous in every component, i.e., the likelihood functions

$$IF_{P;is} : t \rightarrow \mathrm{p}_t(I_i|S_s)$$

are all continuous.

So as in Radner and Stiglitz (1984) with a slightly different interpretation, the information processing activity is real-valued. Diverging from their approach we shall restrict ourselves to decision problems with a finite number of alternative courses of action A_a, environmental states S_s, information signals I_i, and a risk-neutral decision maker. If information processing starts with zero information (completely non-informative information structure), the likelihood $\mathrm{p}_0(I_i|S_s)$ for all signals I_i is independent of the state S_s. Unlike in the Radner-Stiglitz model, however, this is not required.

The signals represent possible outcomes of information processing from a prior perspective. With growing information processing activity t, the information structures attached become more informative, meaning the stochastic indication of the signals toward the states gets increasingly stronger. Eventually but not necessarily, the point of complete information is reached, where the conditional probability $\mathrm{p}_t(S_s|I_i)$ for all states and signals is either 1 or 0. Differing from evaluating processing of information in one single step, the information here is processed continuously. The signals I_i represent prior possibilities for processing results. It should be noted that for the sake of simplicity, the set of possible signals does not depend on the processing activity t.

It must also be kept in mind that the value of the information structures is always calculated from the prior perspective, implying that a decision about an optimal information processing activity derived from the model uses only the information given before processing starts: the decision about whether to stop or to proceed with information processing here is not adaptive in the sense that it uses the signals being processed along the way. This is important since an adaptive strategy might lead to better results when feasible: with the constantly changing base of the decision maker's information while processing, the relevant adaptive information processing function for the remaining processing activity changes steadily.

Since the value of a given information structure (likelihood matrix) $(q(I_i|S_s))_{is}$ is calculated according to (1), and (3) gives the continuous sequence of information structures satisfying (4), the information processing function IPF can be defined according to

$$\mathrm{IPF}_{D;P}(t) = \mathrm{IV}_D(\mathrm{IV}_P(t)) \tag{5}$$

IPF is the 'production function' for the information processing activity, mapping each processing activity t on the prior value of its associated information structure before cost. The information structure function P_t models information processing, i.e., by assumption condition (4) is fulfilled for all $r \leq t$. Hence, it is immediately clear from Blackwell's result that $\text{IPF}_{D;P}$ is monotonously non-decreasing in t.

For every processing activity t and signal I_i, the action with the maximal expected payoff U will be marked with the index $a^*_{i,t}$. In the case of zero information at $t = 0$, clearly $a^* = a^*_{i,0}$ is the optimal action independent of signal I_i without additional information processing. Starting from (1) and (5) using this notation, we obtain

$$\begin{aligned}
\text{IPF}_{D;P} &= \sum_{i=1}^{I} \text{p}_t(I_i) \times \left[\max_{a=1,\ldots,A} \left\{ \sum_{s=1}^{S} \text{p}_t(S_s|I_i) \times U_{as} \right\} \right. \\
&\quad \left. - \sum_{s=1}^{S} \text{p}_t(S_s|I_i) \times U_{a^*s} \right] \\
&= \sum_{i=1}^{I} \text{p}_t(I_i) \times \left[\sum_{s=1}^{S} \text{p}_t(S_s|I_i) \times \left(U_{a^*_{i;t}s} - U_{a^*s} \right) \right] \\
&= \left[\sum_{i=1}^{I} \sum_{s=1}^{S} \text{p}_t(I_i) \times \frac{\text{p}_t(I_i|S_s) \times \text{p}(S_s)}{\text{p}_t(I_i)} \times \left(U_{a^*_{i;t}s} - U_{a^*s} \right) \right] \\
&= \sum_{s=1}^{I} \sum_{s=1}^{S} \text{p}(S_s) \times \left(U_{a^*_{i;t}s} - U_{a^*s} \right) \times \text{p}_t(I_i|S_s) \qquad (6)
\end{aligned}$$

DISCUSSION A closer look at the last term shows an important feature of the information processing function. If for a given signal I_i and activity level t^* there is a single action with maximal expected payoff, which implies that no two or more alternative courses of action have the same expected payoff at (t^*, i), then there will be a neighborhood of t^* for which this action stays optimal for I_i, because IF_P was assumed to be continuous in every component $\text{p}_t(I_i|S_s)$.

Case 1. If this neighborhood can be found at t^* for every signal I_i, then there will also be a neighborhood near t^* in which the actions that are optimal for the respective signals I_i stay the same. In that case, the payoff difference in the last term of (6) is constant within that neighborhood, i.e., it does not depend on t. Since the prior state probabilities $\mathrm{p}(S_s)$ also do not depend on t, the information processing function within that neighborhood is a linear mixture of the likelihood functions' components $\mathrm{p}_t(I_i|S_s)$.

Case 2. If there is a signal I_i for which in t^* at least two actions yield the same expected payoff, then for this signal the optimal action changes at t^*. For these t^* a change of the optimal course of action happens conditionally on a signal I_i.

In Section 4, specific functions IF_P will be analyzed for their associated information processing functions $\mathrm{IPF}_{D;P}$.

4. RESULTS: THE SHAPE OF INFORMATION PROCESSING FUNCTIONS

It is clear that even if all likelihoods $\mathrm{p}_t(S_s|I_i)$ are differentiable in t, the same cannot be said for the information processing function. Formally, this is caused by the maximum function in the value of information.

RESULT 1. If information processing is smooth in the sense that all likelihoods are differentiable everywhere in t, then the information processing function is both continuous and non-decreasing everywhere and is piecewise differentiable with isolated activity levels where it can usually not be differentiated. These t correspond to processing activity levels for which there is a change in the optimal action for certain signals I_i (*'conditional change of action'*). At these non-differentiable levels the function is convex even for concave likelihoods, i.e., the marginal return on its right are larger than those on its left.

The continuity results directly from the continuity of all likelihood functions $\mathrm{p}_t(I_i|S_s)$. In *case 1* of the brief discussion following (6), if all $\mathrm{p}_t(I_i|S_s)$ are differentiable in such a t^*, the same is true for $\mathrm{IPF}_{D;P}$ at this particular t^* and within its neighborhood where the optimal action does not change for any signal I_i. In *case 2* of that discussion, there is a conditional change in the optimal action for at least one signal I_i at t^*. At these t^*, the payoff optimal action and thus the payoff difference in (6) changes for at least one signal. The information processing function usually cannot be differentiated in such t^*, but is clearly still continuous.

The growth in slope follows by looking at the last term in (6): If the function cannot be differentiated at t^* as described, then for at least one signal I_i there are two courses of action yielding the same expected payoff at t^*, and for $t > t^*$ a different action becomes conditionally optimal under I_i in expected payoff. If the decision maker wouldn't change his course of action for $t > t^*$ under this I_i, his information processing function $\text{IPF}_{D;P}$ would sub-optimally stay at the previous non-decreasing path (see (6)). Changing to the conditionally superior action for $t > t^*$ at I_i changes the 'mixing factors' for the likelihood functions $\text{p}_t(I_i|S_s)$ of $\text{IPF}_{D;P}$.[4]

With the model described in Section 3, it is easy to see that starting at $t = 0$ with completely non-informative information and a single best course of action A_{a^*}, a result similar to Radner and Stiglitz applies: at the start, information processing is completely worthless, since the alternative courses of action have to compensate continuously in t for their disadvantage in expected payoff.

RESULT 2. In the model described in Section 3, zero information at $t = 0$, and a single best action A_{a^*}, i.e., a unique action with maximal expected payoff, information processing is completely worthless at the beginning: In this case, there exists an $t^* > 0$ such that for all $t \leq t^*$: $\text{IPF}_{D;P}(t) = 0$.

This result follows directly from (6) and the reasoning at the end of the previous section *(case 1)*. If there is zero information at $t = 0$, then the $\text{p}_0(I_i|S_s)$ are independent of the states S_s, hence $\text{p}_0(S_s|I_i) = \text{p}(S_s)$ for all signals I_i. If there is a unique action with maximal expected payoff, the payoff of the other possible courses of action have to overcome a difference in expected payoff before they can be optimal for certain signals I_i. Since it is assumed that all $\text{p}_t(I_i|S_s)$ change continuously with t, the same is true for $\text{p}_t(S_s|I_i)$. Hence under each signal, A_{a^*} stays optimal for sufficiently small changes of t. Thus taking the minimal one of these sufficiently small changes

TABLE I

Decision problem D and information structure function P_t[5].

D	S_1	S_2	
$p(S_s)$	0,5	0,5	E_U
A_1	220.00	−255.00	−17.50
A_2	150.00	25.00	87.50
A_3	100.00	100.00	100.00
A_4	63.33	130.00	96.67
A_5	−6.67	160.00	76.67
A_6	−366.67	200.00	−83.33
P_t			
$p_t(I_i \mid S_s)$	I_1	$_2$	$p(S_s)$
S_1	$\frac{1}{2}+\frac{1}{2}t^{0.3}$	$\frac{1}{2}-\frac{1}{2}t^{0.3}$	0,5
S_2	$\frac{1}{2}-\frac{1}{2}t^{0.3}$	$\frac{1}{2}+\frac{1}{2}t^{0.3}$	0,5
$p_t(I_i)$	0,5	0,5	

for all signals, we have found a t^* up to which $\mathrm{IPF}_{D;P}$ stays at zero and the processing is of no value (see (6), (1)).

EXAMPLE 1. A risk-neutral decision maker faces a decision problem with six possible courses of action and two environmental states with attached prior probabilities and information structures arising during processing (see Table I).

Figure 1 shows the information processing function and its derivative for this case of differentiable information processing and zero information at activity level $t = 0$.[6] Clearly, the value of perfect information and thus an upper bound for the function is 110. In this special case, $p_t(I_i)$ is independent from t, since $p(S_1) = p(S_2)$.

It seems reasonable to look at information processing functions of simple likelihood functions $p_t(I_i|S_s)$ and, hence, IF_P. The simplest interesting case is a constant change in all likelihoods $p_t(I_i|S_s)$, i.e., the case of linear likelihood functions in each component: we will call information processing linear and the associated learning having a *constant pace* when for all states S_s and all signals I_i there are real parameters a_{is} and b_{is} such that

$$p_t(I_i|S_s) = a_{is} \times t + b_{is} \tag{7}$$

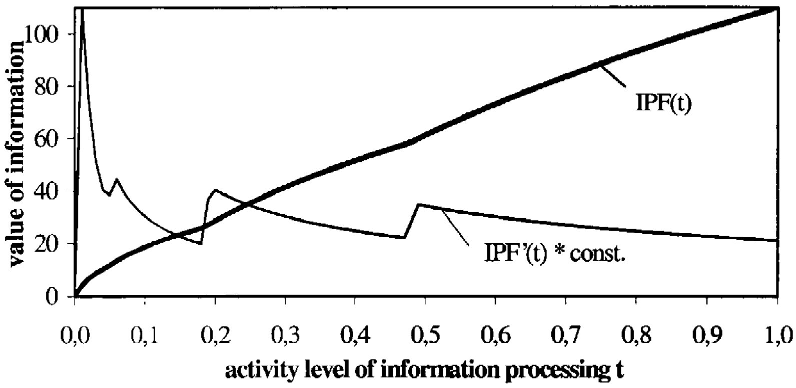

Figure 1. Information processing function for the decision problem and the information structure function in Tables 1 and 2.

If the absolute value of a_{is} is comparatively high for all signals and states, the likelihoods change relatively quickly with t and learning has a high pace and vice versa. It is natural to formulate the pace of learning at t as the sum of the likelihoods' differentials at t in absolute values:

$$\mathrm{LP}_p(I;t) = \sum_{i=1}^{I}\sum_{s=1}^{S}\left|\frac{\partial \mathrm{p}_t(I_i|S_s)}{\partial t}\right| \tag{8}$$

With this definition, linear information processing as defined in (7) has a constant learning pace at all t.

RESULT 3. For linear information processing, the information processing function is piecewise linear between the activity levels where it cannot be differentiated. At these levels, its slope grows.

The piecewise linearity in the intervals between the conditional changes of action (Result 1) follows directly from (6) and the considerations thereafter, together with the linearity of all likelihood functions (see (7)). The convexity at the activity levels where $\mathrm{IPF}_{D;P}$ cannot be differentiated is a special case of Result 1.

EXAMPLE 2. In the decision problem discussed in Example 1 (see Table I), the risk-neutral decision maker now faces an information processing opportunity as described in Table II. The information processing function in this case of linear information processing is shown in Figure 2. At $t = 0$ there is again zero information.

TABLE II

Information structure function for linear information processing.

$p_t(I_i \mid S_s)$	I_1	I_2	$p(S_s)$
S_1	$\frac{1}{2}+\frac{1}{2}t$	$\frac{1}{2}-\frac{1}{2}t$	0.5
S_2	$\frac{1}{2}-\frac{1}{2}t$	$\frac{1}{2}+\frac{1}{2}t$	0.5
$p_t(I_i)$	0.5	0.5	

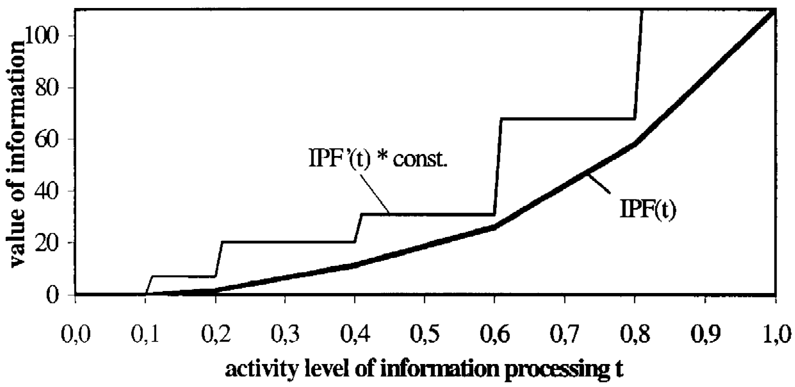

Figure 2. Information processing function for the decision problem D (Table I) and linear information processing (IF in Table II)

5. SUMMARY AND OUTLOOK

It has been shown that increasing marginal returns to information processing, even for single usage, are not restricted to the start of processing activities, but do also apply repeatedly through their course. The value of processing information is a locally linear combination of the associated likelihood functions with the payoff of the respective optimal actions serving as weighting factors. If all likelihood functions are linear, information processing yields piecewise linear expected payoffs with increasing slope at the activity levels with conditional change in optimal action for certain signals.

There are many consequences of convexities in information processing functions. First, for finding the optimal (prior) degree of information processing, the standard marginality criterion for optimality does not apply. Seemingly innocent, general statements like March's (1994, p. 25), that '*rational decision makers can be ex-*

pected to invest in information up to the point at which expected cost equals the marginal expected return' prove to be wrong, since the productivity of information processing has systematically increasing marginal returns: even if marginal returns and (constant) marginal costs equal each other, it might well be useful to continue processing if such a convexity to overcompensate a marginal loss is close at hand. In order to make this argument, one does not have to rely on repeated use of information.

Secondly, there are major implications for issues of coordination, control and optimal size of the firm. Let us name two important ones. Arrow (1975) discussed the positive implications for vertical integration due to uncertainty reduction. Similarly, Radner (1992) discussed whether possible disadvantages of size that manifest in upwardly disproportionate growth in coordination costs can be more than compensated by increasing marginal returns of information.

A third area where important implications can certainly be expected is the organization of knowledge-intense industries and the creation of competitive advantage in these. After analyzing situations and implications of increasing marginal returns, Arthur (1996) provides a management-level overview of such considerations. These and other important implications are largely subject to further research and will not be discussed here.

NOTES

1. Blackwell (1951). Clearly, if such an M exists, Q can never be more informative, since Q can be constructed from P any time when knowing M. More surprising is the necessity for the existence of such an M. Notice that P and Q do not have to have the same number of signals since M is not restricted to having the same number of columns and rows. A more fundamental analysis of information structures appears in McGuire (1972).
2. A formal model is given in (5).
3. The assumptions not discussed are in only important to ensure that the Radner–Stiglitz formal model actually models costly information processing in a reasonable way.
4. Schauenberg (1986) has demonstrated this result in the special case of two environmental states.
5. In interpreting $\frac{1}{2} \cdot t^a$ as the pace of learning, it helps to know that in asymptotic statistics, the quality of statistical results usually improves with a speed of $\sqrt{n}$ on average ($a = 0.5$), where n is the discrete number of independent trials. See Rüschendorf (1988).

6. For visual reasons the derivative is multiplied by a constant factor.

REFERENCES

Arrow, K.J. (1974), *Limits of Organization*. New York: Norton.

Arrow, K.J. (1975), Vertical integration and communication, *Bell Journal of Economics* 6, 173–183.

Arrow, K.J. (1985), Informational structure of the firm, *American Economic Review, Papers and Proceedings* 75, 303–307.

Arthur, W.B. (1996), Increasing returns and the new world of business, *Harvard Business Review* 1996, pp. 100–109.

Blackwell, D. (1951), Equivalent comparison of experiments, *Annals of Mathematical Statistics* 24, 265–272.

March, J.G. (1994), *A Primer on Decision Making.* New York: Free Press.

Marschak, J. (1954), Toward an economic theory of organization and information. In: R.M. Thrall, C.H. Coombs, and R.L. Davis (eds.), *Decision Processes* (pp. 187–220). New York: Wiley.

Marschak, J. (1959), Efficient and viable organization forms. In: Haire, M. (ed.), *Modern Organization Theory* (pp. 307–320). New York.

McGuire, C.B. (1972), Comparison of information structures. In: C.B. McGuire and R. Radner (eds.), *Decision and Organization*, 2nd edn (pp. 101–130). Amsterdam: North-Holland.

Radner, R. (1992), Hierarchy: The economics of managing, *Journal of Economic Literature* 30, 1382–1415.

Radner, R. and Stiglitz, J.E. (1984), A nonconcavity in the value of information. In: M. Boyer and R.E. Kihlstrom (eds.), *Bayesian Models in Economic Theory* (pp. 33–52). Amsterdam: North-Holland.

Rüschendorf, L. (1988), *Asymptotische Statistik.* Stuttgart: B.G. Teubner.

Schauenberg, B. (1986), Der Verlauf von Informationswertfunktionen. In: W. Ballwieser and K.-H. Berger (eds.), *Information und Wirtschaftlichkeit* (pp. 229–251). Wiesbaden: Gabler.

Address for correspondence: Hagen Lindstädt, Department of Strategic Management and Organization, HHL – Leipzig Graduate School of Management, Jahnallee 59, D-04109, Leipzig, Germany
Phone: +49-341-9851676; E-mail: lindstaedt@managem.hhl.de