INFORMATION: ITS MEASUREMENT AND VALUATION

by

Jean-Paul Chavas and Rulon D. Pope*

Invited paper presented at the American Agricultural Economics Association Annual Meetings, Cornell University, August 5-8, 1984.

*The authors are associate professor of agricultural economics, University of Wisconsin and associate professor of economics, Brigham Young University, respectively.

INFORMATION: ITS MEASUREMENT AND VALUATION

I. Introduction

The current "information revolution," exemplified by the increased use of computers and robots in our society, gives a clear indication of the important role that information plays in any social organization. In the past few decades, much progress has been made refining the conceptual tools that can be used in the investigation of information-related issues. However, the analysis of the role of information remains a complex endeavor. For example, it may even be difficult to develop a consensus among scientists on the proper definition of information. For the purpose of this paper, the following taxonomy will be adopted. Data are defined to be the result of an inquiry process (e.g., sampling or experimentation) concerning particular events (e.g., today's hog price in Omaha, corn yield response to fertilizer in Iowa). By using codes (e.g., written or spoken language), the data can be transmitted as signals over space or time through particular communication devices (e.g., radio, telephone, newspapers). These signals generate messages that can be interpreted and used in decision making. For the statistician, the decision may be either to accept or reject a particular hypothesis. Alternatively, for the economist, the decision could be to select and implement a particular production, consumption or investment action. In either case, information will be defined here as the screening, editing and evaluation of data in the context of a particular decision-making process (Caspari).

Any of the functions involving information (data coding, processing and transmission of signals, decision) could be done by men or by machines. The allocation of the tasks between men and machines depends on their relative abilities.[1] For example, humans are very poor transmission channels: Compared to computers, they read or write at a very slow speed and they forget a lot. However, the human language is a very efficient and economical code and the computer is certainly much less complex than the human brain. In general, machines (computers, robots) often appear to have an advantage over humans when performing simple and repetitive tasks. This, coupled with the development of better and cheaper computers seem to be the source of the current "information revolution" which promises to alter significantly business decision making. In this context, it seems important for economists to sharpen the conceptual tools that can be used in the analysis of the role and value of information in management or policy decisions.

The objective of this paper is to briefly discuss the measurement and economic valuation of information. Alternative information concepts found in the literature are reviewed. It is argued that the valuation of information is best analyzed in the context of decision making under uncertainty. A simple model is developed to illustrate how better information tends to improve the decision-making process. Implications of the model for the valuation of information are presented.

II. The Concepts of Information

As different scientists may have different definitions of information, one can find in the literature at least four different concepts and measures of information:[2]

a. The expected information of a probability distribution can be defined as a scalar-valued functional of the probabilities. This is the case of the entropy measure H, defined as minus the expectation of the logarithm of the probabilities.

$$H = -\sum_{i=1}^{n} p_i \ln p_i$$

where p_i is the probability of the i^{th} event (Shannon and Weaver, Theil). The entropy H is a measure of "disorder" in the sense that a larger value of H is associated with more uncertainty. In other words, a gain in information can be measured by a decrease in entropy.[3] This measurement of information has generated considerable enthusiasm in marketing, communications, psychology and other disciplines because of its simplicity. For example, in the context of a multivariate normal distribution with covariance matrix V, then the entropy measure becomes

$$H = \frac{n}{2}(1 + \ln 2\pi) + \frac{1}{2} \ln \det(V)$$

which implies that the amount of information decreases linearly with the logarithm of the determinant of the covariance matrix V. In that context, an experiment that seeks to maximize information also seeks to minimize det(V). This corresponds to D-optimality in the literature on optimal experimental design (Keifer).

Unfortunately, there is in general no relation between the entropy measure for information and the value of information in a particular decision-making process.[4] Nor does there seem to be a relation between the amount of information and the cost of producing it (Marschak). For these reasons, entropy as a measure of information has had little use in economics.

b. Information provided by a sample can also be measured by the information matrix M, defined as the negative of the hessian of the log-likelihood function with respect to the unknown parameters (Fisher).

If the parameters are estimated by a best linear unbiased estimator, then M^{-1} is proportional to the covariance matrix of the estimator. This suggests that the amount of information is closely related to the reliability of the parameter estimates. For statisticians, the information matrix M plays a crucial role in statistical inference. It is also central in the theory of optimal experimental design. For example, common optimality criteria for optimal designs include the minimization of $\det(M^{-1})$ (D-optimality), of $tr[CM^{-1}]$ where C is a given non-negative definite symmetric matrix (L-optimality), or of the maximum eigen value of M^{-1} (E-optimality) (Kiefer). We will see below that the L optimality criterion can provide an attractive approximate measure of the value of information, given an appropriate choice of the C matrix.

c. Information can also be defined as a message which alters tastes or perceptions which are <u>certain</u> (e.g., Auld). This approach has been frequently used in empirical work and in some theoretical analyses (e.g., Hayami and Peterson, Bradford and Kelejian). For example, as a result of a particular message, riskless demand curves may shift. By analogy with the consumer surplus measure, the area between the two curves (before and after the message is received) and above price might then be used to value information. However, there are several weaknesses in this approach. First, it fails to recognize explicitly the role of uncertainty and information in the decision-making process. Secondly, it cannot explain information-gathering activities since, in a riskless world, there is no incentive to learn. Thirdly, it provides an ex-post evaluation of information in the absence of uncertainty. In that sense, it does not measure the ex-ante willingness to pay for information which appears to be the relevant measure in the valuation of information-gathering activities.

d. Finally, information can be defined as a message which alters probabilistic perceptions of random events (e.g., Marschak). This is the approach used in statistical decision theory (Raiffa and Schlaiffer, Lavalle). It seems to have the greatest appeal as a general approach, with wide potential for applications (e.g., Baquet et al.). It is the topic of the next section, where a formal model of information evaluation is developed to illustrate some of its characteristics.

III. A Model of Information

Consider an economic unit (e.g., a firm or a consumer) with a two-period planning horizon: t = 1,2. Denoting its preferences by the Von Neumann-Morgenstern utility function U(.), the objective function of the economic unit is given by

$$\max_{x_1, x_2} E_1 U(y, x_1, x_2, e_1, e_2) \tag{1}$$

where E_t is the expectation operator based on the information available at time t, x_t denotes the decision vector at time t, y is initial wealth, and e_t is a random variable with a given subjective probability distribution reflecting uncertainty (e.g., price uncertainty) in the decision-making process at time t.

If all decisions are made at time t = 1, then (1) corresponds to an open-loop solution where the vectors x_1 and x_2 are chosen at the beginning of the planning horizon, based on the information available at that time.

However, if learning takes place over time, then the decision maker will have more information at time t = 2 than at time t = 1. Given this temporal nature of uncertainty, the decisions can be analyzed through dynamic programming. Using backward induction and starting at the end of the planning horizon, the objective function at time t = 2 is

$$V(x_1, y, e_1) = \underset{x_2}{\text{Max}}\ E_2 U[x_1, x_2, y, e_1, e_2]. \qquad (2a)$$

Expression (2a) indicates that the second period decisions are made conditional on previous decisions x_1. If x_1 does not influence the probability distribution of e_2, it corresponds to a passive learning situation. Alternatively, if x_1 does influence the probability distribution of e_2 by generating a message correlated with e_2, then it corresponds to an active (Bayesian) learning situation where the decision maker uses its own resources to reduce future uncertainty.

Moving back in time, the objective function at time t = 1 is

$$\underset{x_1}{\text{Max}}\ E_1\ V(x_1, y, e_1). \qquad (2b)$$

The dynamic formulation (2) allows for the temporal nature of uncertainty to influence the decision-making process. It corresponds to a feedback solution where learning is explicitly taken into consideration. As such, it will provide the basis for our discussion of information valuation.

The valuation of information relates directly to the comparative performance of decision-making processes given different levels of information. For example, in the context of model (2b), it can be shown that

$$E_1 \underset{x_1}{\text{Max}}\ V\ (x_1, y, e_1) \geq \underset{x_1}{\text{Max}}\ E_1\ V\ (x_1, y, e_1). \qquad (3a)$$

This implies that, on the average, better information tends to improve the decision-making process and make the decision maker better off. It also implies that the value of costless information is non-negative.

Indeed, from (3a), the bid price[5] for obtaining costless information about e_1 before the decision x_1 is made can be defined as the value B_1 which satisfies (Lavalle)

$$E_1 \max_{x_1} V(x_1, y-B_1, e_1) = \max_{x_1} E_1 V(x_1, y, e_1). \quad (3b)$$

The bid price B_1 is the maximum amount of money the decision maker is willing to pay in order to be able to make first-period decisions x_1 knowing e_1.[6] As long as the marginal utility of income is positive ($\partial V/\partial y > 0$), it follows from (3) that $B_1 \geq 0$, i.e., that more costless information can never make the decision maker worse off and can make him better off. A useful illustration of this result can be obtained considering a small increment in information: by taking a Taylor series expansion on both sides of (3b), then, in the neighborhood of perfect information about e_1, the value of information for one unit change in the variance of e_1 can be written as (Hess)

$$B_1 = -\frac{1}{2}\left[\frac{\partial^2 V}{\partial e_1 \partial x_1}\left(\frac{\partial^2 V}{\partial x_1^2}\right)^{-1}\frac{\partial^2 V}{\partial x_1 \partial e_1}\right] \Big/ \frac{\partial V}{\partial y} \quad (4)$$

Given $\partial V/\partial y > 0$ and $\frac{\partial^2 V}{\partial x_1^2}$ being a negative semi-definite matrix under the maximization hypothesis, it follows that the value of a small increment of information in (4) is non-negative as expected. This suggests that, in general, the value of information about e_1 can be approximated by the formula[7]

$$B_1 \cong \left\{-\frac{1}{2}\left[\frac{\partial^2 V}{\partial e_1 \partial x_1}\left(\frac{\partial^2 V}{\partial x_1^2}\right)^{-1}\frac{\partial^2 V}{\partial x_1 \partial e_1}\right] \Big/ \frac{\partial V}{\partial y}\right\} \mathrm{Var}(e_1) \quad (5)$$

which corresponds to the L-optimality criterion in optimal experimental design, with $C = -\frac{1}{2}\frac{\partial^2 V}{\partial e_1 \partial x_1}\left(\frac{\partial^2 V}{\partial x_1^2}\right)^{-1}\frac{\partial^2 V}{\partial x_1 \partial e_1} / \frac{\partial V}{\partial y}$. This measure represents the expected loss associated with suboptimal decisions due to imperfect information about e_1, using a quadratic approximation. Expression (5) is a potentially important result since it provides a formal link between the valuation of information and the analysis of learning activities in empirical research.

Unfortunately, expression (5) gives only a local approximation. In general, this measure will not be globally valid. For example, while (4) or (5) indicates that, locally, more uncertainty (a higher variance of e_1) implies that perfect information is more valuable, Gould and Hess have shown that such a relationship need not hold in general, i.e., that it is possible that an increase in risk be associated with a decrease in the value of information. This indicates that, although both risk and information require the existence of uncertainty, they are basically different economic concepts. While the value of information involves the ex-ante evaluation of an ex-post situation, the economics of risk has been limited mostly to an ex-ante analysis (Arrow, Pratt). Further research exploring the relationships between risk, information and the temporal nature of uncertainty may be promising (e.g., Nachman, Epstein, Jones and Ostroy). For example, it would be helpful to investigate whether a decision maker is less risk averse to a given future risk if he has the opportunity to learn over time and modify his plans as he gets more information.

Arguments similar to (3) can be applied to the second period decision, thus showing that the economic unit is willing to pay a non-negative sum of money for obtaining costless information about e_2 before the x_2 decision is made. Thus, the decision maker has incentives to learn about future uncertainty. In the context of active learning, these incentives can be translated into the use of its own resources x_1 (at time t = 1) to learn about e_2 and reduce future uncertainty.

Similarly, it can be shown that

$$E_1 \max_{x_2} E_2 U(x_1, x_2, y, e_1, e_2) \geq \max_{x_2} E_1 U(x_1, x_2, y, e_1, e_2) \quad (6)$$

This implies that the bid price for obtaining costless information about e_1 before the x_2 decision is made is non-negative, i.e., that the economic unit can benefit from the temporal resolution of uncertainty. Also, it follows from (6) that

$$\max_{x_1} E_1 \max_{x_2} U(x_1, x_2, y, e_1, e_2) \geq \max_{x_1} \max_{x_2} E_1 U(x_1, x_2, y, e_1, e_2) \quad (7)$$

Expression (7) implies that the open-loop solution (1) can never be better than the feedback solution (2). In other words, being able to revise future plans as new information becomes available tends to make the decision maker better off. This suggests that the open-loop models commonly used by agricultural economists may not be appropriate tools of analysis whenever new information has a significant influence on economic decisions.

IV. Some Implications

What are the implications of the discussion of the previous sections? It suggests that information can be considered as an intermediate good: It is the output of an inquiry process which can be either external to the economic unit in the case of passive learning, or internal in the case of active learning. It is also the input into the decision-making process in the sense that it modifies the probability distribution of random parameters of importance in economic decisions. Thus, the supply of information will depend on the cost and characteristics of the inquiry process. Similarly, the demand and value for information will depend on how economic decisions can be improved with better information. Under competition, we can expect the supply to equal the demand for information, optimal information corresponding to the point where the marginal private cost of information equals its marginal private benefits. In the case where information cannot be appropriated or is indivisible, then corrective policies can be developed to try to obtain the "socially optimal" information levels. These arguments hardly seem new for economists.

Perhaps, what makes the economic analysis of information most difficult is a measurement problem. Indeed, because of its subjective nature, it is rather difficult to measure and value information directly. We have argued in the previous section that expression (5) may provide a convenient (approximate) measure of the value of information. Here, we briefly explore some other approaches to the measure of information value, building on the valuation of non-market goods.

One possible approach would be to make use of measurements in related markets. A well-known example of this approach is the travel cost method used in the valuation of outdoor recreation (Clawson). To illustrate, consider the case of active learning where the decision variables x_1 are informational inputs (e.g., managerial time, consulting services, specialized publications, etc.) designed to reduce future uncertainty. Under competition, the vector x_1 is purchased at market price r_1. Then equation (2b) becomes

$$\max_{x_1} E_1 V(y - r_1' x_1, x_1, e_1)$$

which consists in choosing how much to learn about e_2 given the cost of the information-gathering activities x_1. If one of the informational inputs, x_1°, is a necessary input (e.g., managerial time) in the learning process in the sense that learning can take place only if $x_1^o > 0$, then the net value of the information (B) actively gathered by the economic unit is given by the relationship

$$\underset{x_1}{\text{Max}}\ E_1V(y - r_1'x_1 - B, x_1, e_1) = \underset{x_1}{\text{Max}}\ [E_1V(y - r_1'x_1, x_1, e) | x_1^\circ = 0]$$

By differentiating this expression and using the envelope theorem, it follows that

$$B = \int_{\bar{r}_1^\circ}^{\infty} \tilde{x}_1^\circ \, dr_1^\circ \qquad (8)$$

where $\tilde{x}_1^\circ$ is the compensated demand function for the necessary input x_1°, and $\bar{r}_1^\circ$ denotes the current market price of x_1°. Therefore, the area under a compensated informational input demand function and above price can provide an exact measure of information value. As in the consumer surplus literature, one can raise the issue of whether compensated and uncompensated functions have similar price slopes, i.e., whether B can be measured approximately from (8) by replacing $\tilde{x}_1^\circ$ by the observable (uncompensated) demand for x_1° (See Pope et al.). In any case, from (8), the demand for necessary informational inputs can provide an appropriate measure for information valuation. As such, it may be of considerable interest in the empirical evaluation of active learning by economic units.

A possible alternative approach is the contingent valuation method, which typically relies on bidding games (Randall et al., Brookshire and Crocker). For example, the respondents could be asked questions in an attempt to determine their willingness to pay for a particular message or a particular experiment. This approach would be subject to the concerns that our economic models may not be able to explain accurately economic behavior under uncertainty (Kahneman and Tversky, Heiner), and that contingent valuation may be biased (Bishop and Heberlein). However, too little work has been done on this subject to properly evaluate the merits of the contingent valuation approach in the analysis of information value. Further research comparing the relative merits of the alternative approaches just discussed seems greatly needed.

V. Conclusion

There seems to be no question that information is valuable and can play a crucial role in the decisions of all economic agents. For example, the market efficiency hypothesis relies on the way information is processed in a particular market. It seems fair to say that, although some progress has been made in the economics of information over the last few decades, much remains to be done.

We have reviewed what we feel is the proper conceptual way to value information (e.g., 3b). Yet, a number of issues have surfaced which generally require further empirical microeconomic work to resolve:

(a) Under what circumstances is a scalar measure of information such as entropy useful in information valuation?

(b) How do economic agents process information? Are Bayesian revisions descriptive of behavior? How prevalent is active learning and how serious are modeling errors when open loop (as opposed to closed loop) paradigms are employed?

(c) How descriptive is the approximation to information valuation given in (4) and (5) for both small and large projects?

(d) How do valuation results obtained from direct elicitation compare with estimates derived using indirect methods and market data?

Since there is so much inherent uncertainty in the agricultural sector, we have an excellent setting to gain deeper insights regarding these questions and their answers.

Footnotes

[1]Note that similar arguments can be made concerning the allocation of tasks among humans, i.e., the problem of organization in terms of who does what within a particular institution (e.g., a corporation). In this context, the concept of human capital is closely related to the ability of processing information: Persons with better human capital (because of a higher level of education or more experience) may better process information in decision making and, for that reason, be better managers.

[2]An additional possible measure of information involves the use of a sufficient statistic in the comparison of information sets (Blackwell). Generally, this is associated with approach d. below.

[3]In the analysis of noisy messages providing imperfect information, the entropy concept can be used as follows. Denote by p_i the prior probabilities that are modified into the posterior probabilities $p_{i|j}$ given a message j. If q_j is the unconditional probability of the j^{th} message, then a measure of the expected information of the messages is given by the rate of transmission (Shannon and Weaver, Arrow).

$$R = \sum_j q_j \sum_i p_{i|j} \ln (p_{i|j}/p_i) .$$

$$= H (p_i) - \sum_j q_j H (p_{i|j})$$

[4]One exception is the case of a competitive pure exchange economy with a logarithmic utility function. In this context, Arrow has shown that the value of information (defined as the difference in the utilities achievable with and without the information) is equal to the entropy measure H under perfect information, or to the rate of transmission R under noisy messages (see footnote 3). However, since this

measure is expressed in terms of utils, this case would appear attractive in economic analysis only if utilities are additive and commensurable with dollars.

[5]The reservation price or selling price of information could be defined in a similar way (see Lavalle).

[6]Note that similar results hold as well for partial information (see Lavalle).

[7]If e_1 is a vector, then the value of information B_1 would be given by the trace of the matrix on the right hand side of (5).

References

Arrow, K. J. Essays in the Theory of Risk Bearing. Markham Pub. Co., Chicago, 1971.

Auld, D. A. "Imperfect Knowledge and the New Theory of Demand." Journal of Political Economy 80 (1972):1287-1294.

Baquet, A. E., A. N. Halter and F. S. Conklin. "The Value of Frost Forecasting: A Bayesian Appraisal." American Journal of Agricultural Economics 58 (1976):511-520.

Bishop, R. C. and T. A. Heberlein. "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?" American Journal of Agricultural Economics 61 (1979):926-930.

Blackwell, D. "Equivalent Comparison of Experiments." Annals of Mathematical Statistics 24 (1953):265-272.

Bradford, D. F. and H. H. Kelejian. "The Value of Information for Crop Forecasting with Bayesian Speculators: Theory and Empirical Results." The Bell Journal of Economics 9 (1978):123-158.

Brookshire, D. S. and T. D. Crocker. "The Advantage of Contingent Valuation Methods for Benefit-Cost Analysis." Public Choice 36 (1981):235-252.

Caspari, J. A. "Fundamental Concepts of Information Theory." Managerial Accounting 49 (1968):8-10.

Clawson, Marion. Methods of Measuring the Demand for and Value of Outdoor Recreation. Resources for the Future, Washington, DC, 1959.

Epstein, L. J. "Decision-Making and the Temporal Resolution of Uncertainty." International Economic Review 21 (1980):269-283.

Fisher, R. A. "On the Mathematical Foundations of Theoretical Statistics." Transactions of the Royal Philosophical Society London 222 (1922):309-368.

Gould, J. P. "Risk, Stochastic Preference, and the Value of Information." Journal of Economic Theory 8 (1974):64-84.

Hayami, Y. and W. Peterson. "Social Returns to Public Information Services: Statistical Reporting of U.S. Farm Commodities." American Economic Review 62 (1972):119-130.

Heiner, R. A. "The Origin of Predictable Behavior." American Economic Review 73 (1983):560-595.

Hess, James. "Risk and the Gain from Information." Journal of Economic Theory 27 (1982):231-238.

Jones, R. A. and J. M. Ostroy. "Flexibility and Uncertainty." Review of Economic Studies 51 (1984):13-32.

Kahneman, D. and A. Tversky. "Prospect Theory: An Analysis of Decisions Under Risk." Econometrica 47 (1979):263-291.

Kiefer, J. "General Equivalence Theory for Optimum Designs." The Annals of Statistics 2 (1974):849-879.

Lavalle, I. H. Fundamentals of Decision Analysis. Holt, Rinehart and Winston, New York, 1978.

Marschak, J. "Economics of Inquiring, Communicating, Deciding." American Economic Review 58 (1968):1-18.

Nachman, D. C. "Risk Aversion, Impatience and Optimal Timing Decisions." Journal of Economic Theory 11 (1975):196-246.

Pope, R. D., J. P. Chavas and R. E. Just. "Economic Welfare Evaluation for Producers Under Uncertainty." American Journal of Agricultural Economics 65 (1983):98-107.

Pratt, J. W. "Risk Aversion in the Small and in the Large." _Econometrica_ 32 (1964):122-136.

Raiffa, H. and R. Schlaiffer. _Applied Statistical Decision Theory_. Division of Research, Harvard, Massachusetts, 1961.

Randall, A., B. C. Ives and E. Eastman. "Bidding Games for Valuation of Aesthetic Environmental Improvements." _Journal of Environmental Economics and Management_ 1 (1974):132-149.

Shannon, C. and W. Weaver. _The Mathematical Theory of Communications_. University of Illinois Press, Urbana, 1949.

Theil, H. _Economics and Information Theory_. North Holland Publishing Co., Amsterdam, 1967.