ECO 317 – Economics of Uncertainty – Fall Term 2009
Notes for lectures
14. Economics of Information – Introduction

1. Nature of Information – Taxonomies

If one thinks of the economics of information by analogy with the economics of education, health, labor, or whatever, a natural approach would be to develop theories of the demand for information (the value of information), and of the supply of information (the technology and cost of producing information), and then go on to consider a market for information, the socially optimal amount of information, possible market failures, and so on. This is indeed a valid approach to the topic, but it is not directly material to this course. Our concern will not be the market for information per se, but how the limited availability of information in markets for other goods and services, and particularly the asymmetric availability of information to different participants in these markets, affects their outcomes. That can of course be a useful input to the study of the performance of markets for information itself.

However, it is useful to begin with a brief taxonomy of various dimensions and interpretations of information:

[1] One can distinguish between the stock of information (accumulated knowledge) and flow of new information (increments to the knowledge). There can be separate economic activities and transactions related to each of these. Thus the writing, printing and publishing of encyclopedias has to do with the stock; that of research monographs has to do with the flow.

[2] The activities of communication of information can be considered, along the dimensions of bilateral or multilateral communication to specific audiences or broadcasting, and also along the stock-flow dimensions (basic teaching versus research seminars). Communication can also have a strategic dimension – deliberate misrepresentation or manipulation of the truth, and we will consider some of this.
For more on such matters, you can read pp. 167-170 of the Hirshleifer-Riley book on the optional reading list.

2. Value of Improved Decision-Making

In the context of uncertainty, the most important aspect of information is the conditioning and Bayesian updating of the probabilities of various possible outcomes or states of the world. This implies a value of information, namely the increase in utility or payoff that is made possible by the ability to make decisions in the light of a better knowledge of the probabilities.

Value of Full Information

As a simple start, suppose there are S states of the world, $s = 1, 2, \ldots S$, with probabilities forming the vector $\mathbf{P} = (p_s)$. You can take an action $a \in A$. The outcomes depend on both the action and the state. You have a utility-of-consequences function $u(a, s)$ and the expected utility function

$$EU(\mathbf{P}, a) = \sum_{s=1}^{S} p_s \; u(a, s) \,. \tag{1}$$

Contrast two circumstances: [1] You must choose a ex ante, without knowing which state of the world is going to materialize. [2] You are allowed to choose a ex post, after the uncertainty is resolved and you know the state.

In [1], suppose your optimum choice is $a^{[1]}$. In [2], suppose your optimum choice in state s is $a^{[2]}(s)$. Since this maximizes $u(a,s)$, obviously

$$u(a^{[2]}(s), s) \geq u(a^{[1]}, s)\,.$$

There may be an equality in one or a few states if the optimal ex ante choice happens to be optimal ex post in that state or those states, but that may not be the case in any state so the inequality is strict in every state. But except in trivial situations where the same action is optimal in all states of positive probability, we will have

$$\sum_{s=1}^{S} p_s \; u(a^{[2]}(s), s) > \sum_{s=1}^{S} p_s \; u(a^{[1]}, s)\,. \tag{2}$$

The difference is often called the value of information. This is OK for brevity, but it is more accurate to call it the value of *information service*. That is, suppose someone knows or is going to know the true state before you do. Before the uncertainty is resolved, you sign a contract with this person; he promises to tell you the information and then you make the decision that is optimal for that state. Without such a contract, you would have to make the decision knowing only the probabilities, not the actual state. How much would you be willing to pay for the contract?

In utility terms, it is just the difference between the left and right hand sides in (**??**). In money terms, we have to define the monetary consequence function say $C(a,s)$ and then take the utility $u(C(a,s))$. The maximum fee you would be willing to pay is then the F that solves the equation

$$\sum_{s=1}^{S} p_s \; u(C(a^{[2]}(s), s)\,) - F) = \sum_{s=1}^{S} p_s \; u(C(a^{[1]}, s)\,)\,.$$

Value of Partial Information

More typically, one gets partial or interim information but not a full resolution of uncertainty. This can be modeled as follows. The possible items of interim information are the different probability vectors $\mathbf{P}^m = (p_s^m)$ for $m = 1, 2, \ldots M$. The probability of getting interim information m is q_m. So before you get the interim information, the probability of state s is

$$\sum_{m=1}^{M} q_m \; p_s^m\,.$$

The vector of these probabilities is just the vector sum

$$\mathbf{Q} \equiv \sum_{m=1}^{M} q_m \; \mathbf{P}^m\,.$$

If you must make your decision a before you get the interim information, you will make it optimally to maximize (**??**) for this probability vector. Write the resulting expected utility as

$$V(\mathbf{Q}) \equiv \max_{a \in A} \; EU(\mathbf{Q}, a) \,. \tag{3}$$

If you can make the decision after you get the interim information, you will choose for each m the optimal a^m to get

$$V(\mathbf{P}^m) = \max_{a^m \in A} \; EU(\mathbf{P}^m, a^m)$$

and the expected value of this ex ante is

$$\sum_{m=1}^{M} \; q_m \; V(\mathbf{P}^m) \,. \tag{4}$$

Which is bigger: (**??**) or (**??**)? To answer this, note that the V function defined here is exactly the utility function representing *induced preferences* that we came across when we studied critiques of expected utility theory (Lecture material handout No. 10, p. 11). There I argued that the function is convex. Using that property, we have

$$V(\mathbf{Q}) = V\left(\sum_{m=1}^{M} \; q_m \; \mathbf{P}^m \right) \le \sum_{m=1}^{M} \; q_m \; V(\mathbf{P}^m) \,.$$

Therefore making the decision in the light of the interim information is preferable. The right hand side minus the left hand side is the value of the interim information in utility terms.

When discussing induced preferences, I gave a heuristic argument for the convexity: for each fixed action a, the expected utility in (**??**) is a linear function of the probabilities, so its graph in $(S+1)$ dimensional space is linear (the S probabilities in the unit simple of the S-dimensional space constituting the independent variables and the expected utility being the vertical axis or the dependent variable). When the action is chosen optimally for each $\mathbf{P}$, the V function is the upper envelope of all these linear functions; therefore it is convex. The textbook gives the same heuristic argument (bottom of p. 127). But for completeness or rigor, here is a formal proof.

Consider any two probability vectors $\mathbf{P}^1$ and $\mathbf{P}^2$, and any $\theta \in [0,1]$. Form the averaged vector

$$\mathbf{P}^3 = \theta \; \mathbf{P}^1 + (1-\theta) \; \mathbf{P}^2 \,.$$

Suppose the action a^* is optimal for $\mathbf{P}^3$. It doesn't have to be optimal for either $\mathbf{P}^1$ or $\mathbf{P}^2$. Therefore

$$\begin{aligned} V(\mathbf{P}^3) &= EU(\mathbf{P}^3, a^*) = \sum_{s=1}^{S} \; p_s^3 \; u(a^*, s) \\ &= \sum_{s=1}^{S} \; [\theta \; p_s^1 + (1-\theta) \; p_s^2] \; u(a^*, s) \\ &= \theta \; \sum_{s=1}^{S} \; p_s^1 \; u(a^*, s) + (1-\theta) \; \sum_{s=1}^{S} \; p_s^2 \; u(a^*, s) \end{aligned}$$

$$
\begin{aligned}
&= \theta\ EU(\mathbf{P}^1, a^*) + (1-\theta)\ EU(\mathbf{P}^2, a^*) \\
&\leq \theta\ V(\mathbf{P}^1) + \theta\ V(\mathbf{P}^2)
\end{aligned}
$$

Observe that in getting the $\leq$ in the last line we must use the definition of the V function as the maximum, and the fact that θ and $(1-\theta)$ are both non-negative.

3. Option Values

The value of information can be interpreted alternatively and equivalently as the value of the ability to postpone the decision until the uncertainty is resolved (fully or partially as appropriate in the context of the information). Suppose the decision is a simple binary one, of whether to invest I in a project or an asset. In state s, this will yield return $R(s)$. Arrange or label the states so that $R(s)$ is an increasing function. If you make the investment, then in state s your net payoff or profit will be $R(s) - I$.

For simplicity of analysis, suppose that s is a continuous variable with a probability density function $f(s)$ over the support $[s_{min}, s_{max}]$. The problem is non-trivial only if $R(s_{min}) < I < R(s_{max})$, so suppose this is the case. Suppose also that in the absence of information about the actual state, it is optimal for you to invest, i.e.

$$
\int_{s_{min}}^{s_{max}} R(s)\, f(s)\, ds > I\,.
$$

Your expected profit from investing is then

$$
\int_{s_{min}}^{s_{max}} R(s)\, f(s)\, ds - I = \int_{s_{min}}^{s_{max}} [\, R(s) - I\,]\, f(s)\, ds\,.
$$

You would prefer to make the investment decision after finding out the state. Let s^* be defined by $R(s^*) = I$. So you would like to invest if $s > s^*$ and not invest if $s < s^*$. Looking ahead to this from the current perspective when you do not yet know the state, your expected profit is then

$$
\int_{s^*}^{s_{max}} [\, R(s) - I\,]\, f(s)\, ds\,.
$$

Therefore the value of the ability or *option* to postpone the decision is

$$
\begin{aligned}
\int_{s^*}^{s_{max}} [\, R(s) - I\,]\, f(s)\, ds - \int_{s_{min}}^{s_{max}} [\, R(s) - I\,]\, f(s)\, ds &= -\int_{s_{min}}^{s^*} [\, R(s) - I\,]\, f(s)\, ds \\
&= \int_{s_{min}}^{s^*} [\, I - R(s)\,]\, f(s)\, ds > 0
\end{aligned}
$$

since $I > R(s)$ for $s < s^*$.

This is *Bernanke's bad news principle*: the value of the ability to get better information before making an irreversible decision is governed by the distribution of the bad outcomes, namely those where you would prefer not to invest. The reason is that these are the losses that the ability to wait helps you avoid. (Bernanke, QJE 1983.)

Instead of a physical investment, this could be a financial investment: a call option on a stock that gives you the right, but not the obligation, to buy it for a preset price I. If the actual price $R(s)$ of the stock in state s turns out to be higher, you will exercise your option and make a profit $R(s) - I$. If not, you will let your option lapse unexercised, and make zero. What will you pay for such an option right now? Owning the option is equivalent to owning $R(s) - I$ Arrow-Debreu securities in states $s > s^*$. How do we find the value of this in today's market? If there are complete financial markets, we know (directly, or as an approriate linear combination of prices of available securities) the equilibrium price of a state-s, say $p(s)$. Then the option has value

$$\int_{s^*}^{s_{max}} [\, R(s) - I \,]\, p(s)\, ds \,.$$

If markets are incomplete, the prices $p(s)$ may involve some individual-specific (subjective) elements. Also, in practice these formulas are more complicated because of [1] time-discounting, [2] foregone flow payoffs (such as dividends on a stock) while you wait. These matters are covered in specialized and more advanced courses in finance.

4. Asymmetric Information – Taxonomies

Our main focus for the next two weeks will be on economic transactions where the parties do not have equal information. There are various aspects of this.

First, the kinds of transactions that are feasible depend on the extent to which information is public or private. There are three main categories:

[1] Verifiable information – This can be demonstrated to outsiders according to an accepted standard of evidence or proof. Contracts or transactions where what the participants are supposed to do is conditional on verifiable information can be formal, enforceable in a court.

[2] Observable information – This can be seen by all parties to the transaction, but cannot be proved to outsiders. Contracts based on such information have to be self-enforcing, for example based on equilibria of repeated games in ongoing relationships.

[3] Private information – Known only to one party in the transaction, not observable by others. Then only those contracts or transactions that are compatible with the private incentive of this party are feasible.

Next, we distinguish asymmetry that exists before the contract is written or the transaction is entered into, and asymmetry that arises during the transaction. Examples of the former are: each person knows his own innate productivity or risk before entering into an employment or insurance contract, but the employer or insurer does not. These are situations of *adverse selection*. Examples of the latter are: unobservable actions of one party, e.g. the effort or care exercised in employment or mitigation of risk. These are situations of *moral hazard*. There is a related post-contract asymmetry where one party observes something without being able to control it directly, e.g. the size of the loss may be less perfectly observed by the insurance company than by the insured, creating incentives for misrepresentation. This is *costly state verification* (inability of one side to distinguish between states with different losses); we will not consider this much or at all.

Our typical transaction will involve a contract between a *principal* and an *agent*. We usually assume that the agent has the information advantage and the principal proposes the contract, but there are other possibilities. The principal will choose the contract bearing in mind the informational advantage and the incentives of the agent. The contract will also be affected by the agent's opportunities outside of the relationship with this principal, and by the competition that may exist among several potential agents or several potential principals. We will consider several examples, involving different degrees of competition, one-time versus repeated relationships, one principal transacting with many agents and vice versa, hierarchies of agencies, different ways in which adverse selection and/or moral hazard manifest themselves, and so on.

While there are information asymmetries, we suppose that at some level the structure of the interaction is *common knowledge* among the participants. Thus a potential employee may have high or low productivity, and the employer does not know which, the probabilities of it being high or low are known to the employer. Not only that; the employee knows that the employer knows these probabilities, the employer knows that the employee knows that the employer knows the probabilities, and so on *ad infinitum*. This makes game-theoretic modeling of the situations possible.