ORIGINAL ARTICLE

The Physical Foundation of the Mind

Jing Chen

Abstract

The mind of animals, just like the bodies of animals, is shaped by natural selection to best adapt to physical laws. Specifically, the mind is evolved to search for natural resources at low cost. Since entropy provides a universal measure of resources, it is inevitable that information, which we collect for our survival, is represented by the entropy function mathematically. When observing the problems of the mind from the lens of physical laws, many complex and confusing problems become simple and clear. As an example, we present a mathematical theory on the value of judgment in this paper. It is generalized from the entropy theory of information when subjective assessment of the probability distribution of a random event differs from the objective probability distribution. The formula of the value of judgment defined in this theory bridges the chasm between the concept of information and cognitive bias. We prove that the first order approximation of the expected rate of return of the portfolio constructed from a certain judgment is equal to the value of the judgment defined in this theory. This theory on the value of judgment, together with other recently developed theories, shows that research on mind, just like research on matter, can be carried out in an analytical, quantitative and systematic way from the foundation of physical laws.

Key Words: mind, physical laws, judgment, information, entropy, bias, investment return

NeuroQuantology 2008; 3:222-233

Introduction

To gain a deep understanding of human mind, we must first understand why and how the mind is evolved. Most would agree that the streamlined shape of fish is an adaptation of aquatic environment to reduce the cost of swimming. Similarly, the mind of human beings, as well as other animals, is evolved to find resources with low cost. While the forms of natural resources are diverse, most natural resources can be understood from a unifying principle. A system has a tendency to move from a less probable state to a more probable state. This tendency of directional movement is the source of useful energy that drives, among other things, the living organisms. Intuitively, natural resources are something that is of low probability, or scarce. The dissipation of a system from a low probability state into a high probability state provides the energy we use.

The measure of probability of a system is called entropy in physics. In a formal language, systems move from low entropy state to high entropy state. Since all living organisms need to extract low entropy from the environment for survival, it is inevitable that the human mind is evolved to identify entropy easily. This explains why information, which we collect for our

Corresponding author: Jing Chen
Address: School of Business, University of Northern British Columbia Prince George, BC, Canada V2N 4Z9
Phone: + 1-250-960-6480
e-mail: chenj@unbc.ca

survival, is represented mathematically by the entropy function. Human mind, including innate mathematical ability, is evolved to reduce the cost of identify natural resources. When observing the problems about the mind from the lens of physical laws, many complex and confusing problems become simple and clear. For example, the entropy theory offers a unified understanding of the common patterns of the human mind documented in literature (Chen, 2003, 2005). In this work, we will show how the intuition from statistical mechanics helps us develop a mathematical theory on the value of judgment.

Superior investment returns are often attributed to superior information processing capabilities of investors. Inferior investment returns are often attributed to inferior cognitive capabilities of investors. While investment return can be quantified, there does not exist a quantitative measure of judgment. We will develop a mathematical theory of judgment. It provides a quantitative measure on the value of judgment.

This theory of judgment is generalized from the entropy theory of information. From information theory, the value of information of a particular state is the log function of the probability of that state. The value of information of a random event is the average of information value of each possible state of the random event, which is the entropy of the random event. In decision making process, the default position is usually taken to be the maximum entropy equilibrium state. Since no additional information is required to determine the equilibrium state, the value of information from the decision making perspective is the entropy of the random event minus the maximum entropy.

The subjective judgment of the probability of a state by a person may differ from its objective probability. We propose that, the value of judgment of a particular state is the log function of the subjective assessment of the probability of that state. The value of judgment of a random event is the average of value of judgment of each state, weighted by the objective probability distribution of the random event. We will call this value subjective entropy. Similar to value of information, we define value of judgment from the decision making perspective as the subjective entropy minus the maximum entropy.

Mathematically, suppose $\{p_1, \ldots p_n\}$ is the objective probability distribution of a random event and $\{q_1, \ldots q_n\}$ is the subjective judgment of the probability distribution by someone. The value of information is

$$\sum_{j=1}^{n} p_j \ln(p_j) + \ln n$$

While the value of judgment is

$$\sum_{j=1}^{n} p_j \ln(q_j) + \ln n$$

To evaluate the validity of the new theory, we will compare the value of a judgment calculated from this theory with the expected rates of return of the portfolio constructed from the same judgment. Investment decisions are made according to investors' judgment about stocks. For example, if an investor believes one stock will significantly outperform the general market, he will put a significant portion of his investment fund into that stock. To quantify the relation between judgment and investment decision, we will consider a simple portfolio with only two assets: a risk free asset and a risky asset. Based on the subjective assessment of the return distribution of the risky asset, an investor can determine the optimal portion of the risky asset in the portfolio and calculate the expected rate of return of the portfolio. We prove that the first order approximation of the expected rate of return of the portfolios constructed from a judgment is equal to the value of the same judgment. This indicates that the value of judgment defined in this theory is built on a solid foundation.

There are some other general properties from the mathematical theory on the value of judgment. First, the value of judgment is always lower than or equal to the value of information with the same objective probability distribution. This means that the value of information is the upper bound for the value of judgment. A similar idea has been developed in evolutionary biology (Bergstrom and Lachmann, 2004). The value of judgment is equal to the value of information only when the subjective assessment of the probability distribution is identical to the objective probability distribution. Therefore, the concept of judgment is a generalization from the concept of information when a person may not have precise estimation of a random event, which is the case in most decision making processes. The theory of judgment broadens the scope of applications with respect to the theory of information. In particular, the values of information are always nonnegative while the values of judgment can be either positive or negative. This means that an active investment strategy may earn either positive or negative returns with respect to a passive investment strategy.

Second, the value of judgment has its parallel in the works of the coding theory. Sims (2003) described knowledge as a process of coding. A major purpose of coding is to reduce the average length of codes from knowledge about the probability distribution of events. In general, more frequent events are represented by shorter codes. This coding process reduces the average length of the message. Hence, the value of a coding is the reduction of average coding length from the generic coding (Cover and Thomas, 2006). The value of judgment shares the same mathematical formula as the value of coding.

The average length of an actual coding is often longer than the theoretically shortest length possible. In coding theory, the difference between the actual coding and theoretically optimal coding is defined by relative entropy (Cover and Thomas, 2006), which is the difference between the value of information and the value of judgment. Hence relative entropy can also be used to define the measure of bias in the theory of judgment. The parallel between the theory of judgment and the theory of coding, which is a well established subject, is a further indication on the validity of the theory of judgment.

Both the value of judgment and the measure of bias are mathematical functions whose only variables are objective probability distributions and subjective probability distributions, i.e., our judgment, of random events. This makes it very easy to systematically investigate the properties of the theory of judgment.

Third, under certain conditions, a judgment that is more biased turns out to be more valuable than a less biased judgment. Intuitively speaking, an investor who is modestly favorable to a stock which turns out to earn very high rate of return will perform better than an investor who is modestly favorable to a stock which turns out to earn moderately high rate of return. This shows that value and bias of judgment are two distinct concepts. It will help clarify discussion in behavioral literature, which often identifies bias with low value of judgment.

There are several other distinctive properties of this mathematical theory of judgment, which we will derive later in this paper. The theory of judgment bridges the chasm between the concept of information and cognitive bias. This will help provide a common framework for behavioral and informational perspectives in understanding financial market. A detailed discussion will be provided in a companion paper (Chen, 2008).

This theory comes out of the long term effort to understand the Gibbs inequality, a result in statistical mechanics from more than a hundred years ago. It is also inspired by several recent works. Sims (2003) described knowledge as a process of coding. I attempt to learn more about the coding theory, from which the theory of judgment is developed. Bergstrom and Lachmann (2004) discussed the relation between the entropy measure of information and the measure of fitness, which helped me to seek the relation between the measure of judgment and the measure of investment returns. Qian (2001) discussed the usefulness of relative entropy in formulating non-equilibrium statistical mechanics. This deepens my appreciation on the close connection between entropy as a mathematical theory and as a physical theory.

The remainder of the paper is structured as follows. Section I presents the generalized entropy theory of information. It offers a deep connection among information, entropy and human mind. In Section II, we will give a detailed presentation of the mathematical theory on the value of judgment. In Section III, we discuss the relation between value of a judgment and the expected rate of return of the portfolio constructed from the same judgment. Section IV concludes.

I. Resource, Entropy, Information, and the Evolution of Mind

The entropy theory of information can be traced to Maxwell and Boltzmann, the pioneers in statistical mechanics. In a famous thought experiment called Maxwell's Demon, Maxwell (1871) reasoned that if information is costless, then the entropy of the universe can be decreased, which violates the second law of thermodynamics. Therefore the physical cost of obtaining information must be at least equal to its entropy value. Boltzmann also linked the increase of entropy in a system to the loss of information (Isihara, 1971). Shannon (1948) developed the mathematical theory of communication by formally defining information as the entropy function. Shortly after Shannon's work, Weaver commented,

> "Thus when one meets the concept of entropy in communication theory, he has a right to be rather excited --- a right to suspect that one has hold of something that may turn out to be basic and important" (Shannon and Weaver, 1949, p. 13).

The formal link between entropy and information sped up the acceptance of new information theory. However, when the information theory became firmly established, Shannon turned uncomfortable with the early and more speculative attempts to apply the entropy theory of information to broader areas.

> Workers in other fields should realize that that the basic results of the subject are aimed at a very specific direction, a direction that is not necessarily relevant to such fields as psychology, economics, and other social sciences. Indeed, the hard core of information theory is essentially, a branch of mathematics, a strictly deductive system. (Shannon, 1956)

Despite many fruitful applications of information theory to problems in other areas, such as Jaynes (1957), this view remains orthodox today, as indicated by the following comment:

> The efforts of physicists to link information theory more closely to statistical physics were less successful. It is true that there are mathematical similarities, and it is true that cross pollination has occurred over the years. However, the problem areas being modeled by these theories are very different, so it is likely that the coupling remains limited.
>
> In the early years after 1948, many people, particularly those in the softer sciences, were entranced by the hope of using information theory to bring some mathematical structure into their own fields. In many cases, these people did not realize the extent to which the definition of information was designed to help the communication engineer send messages rather than to help people understand the meaning of messages. In some cases, extreme claims were made about the applicability of information theory, thus embarrassing serious workers in the field. (Gallager, 2001, p. 2694)

However, the dissonance between the entropy function as a mathematical representation of information and the practical value of information has long puzzled many people. Recent works have shown that our intuitive concept of information coincides with the mathematical definition of information as entropy (Adami, 2004; Bergstrom and Lachmann, 2004; Qian, 2001). In the following we will provide a more formal argument. If a decision making process is truly important and is needed again and again in life, it is highly economical that quantitative modules evolve in the mind to expedite the process. For example, predators need to routinely assess their distance from the prey, the geometry of the terrain, the speed differential between itself and the prey, the energy cost of chasing down its prey, the probability of success of each chase and the amount of energy it can obtain from prey to determine whether, when and where to initiate a chase. There are many other sophisticated functions, such as navigation by migrating birds over long distances, which need sophisticated mathematical capabilities. Many animals need to make precise calculations of many of these quantitative problems many times in life. To reduce the cost of estimation, mathematical models must have evolved in their mind so that many decision making processes are simplified into parameter estimation and numerical computation. It is highly likely that, if a function is very important for the survival of the animal, in the process of evolution, this function will be genetically assimilated.

From the second law of thermodynamics, a system has a tendency to move from a state of low entropy to a state of high entropy. This tendency of directional movement is the source of useful energy that drives, among other things, the living organisms. Mathematically, the increase of entropy is equivalent to the decrease of useful energy, which is defined as free energy in physics (p. 85, Avery, 2003). Since all living organisms need to extract low entropy from the environment to compensate the continuous dissipation, and entropy is the only mathematical function to measure scarcity of resources (Chen, 2005), it is inevitable that information, which we collect for our survival, is largely about entropy. It is not a mere coincidence that our intuitive concept of information and the mathematical definition of information as entropy largely overlap. The intuitive concept is really a simplified evaluation of a mathematical computation.

If Shannon's entropy theory of information is purely a mathematical theory with little connection to the physical laws, it would be a miracle that information defined as entropy turns out to have the magic properties that handle technical problems in communication so well. However, once mathematical theories are thought to be a natural part of our evolutionary legacy, it would be natural for the entropy theory of information to possess this property.

After the entropy theory of information was developed in 1948, its technique has been applied to many different problems in economics and finance. However, the standard economic theory of information, represented by Grossman and Stiglitz (1980), was not built on the foundation of entropy theory. Economists often feel that "the well-known Shannon measure which has been so useful in communications engineering is not in general appropriate for economic analysis because it gives no weight to the value of the information. If beforehand a large manufacturer regards it as equally likely whether the price of his product will go up or down, then learning which is true conveys no more information, in the Shannon sense, than observing the toss of a fair coin." (Arrow, 1983) The Shannon measure actually carries weight of information. For example, *N* symbols with identical Shannon measure carry *N* times more information than a single symbol (Shannon, 1948). Similarly, the value of the information about the future price is higher to a large manufacturer than to a small manufacturer, other things being equal.

Recently, entropy theory of information has been successfully applied to understand more economic phenomena. Sims (2003) utilized the capacity constraint in information processing to explain price stickiness. This concept has been applied to better understand some empirical patterns in the asset market (Peng, 2005; Peng and Xiong, 2006). The significance of this idea is that it introduces the physical constraint of information processing into economic literature. Sims noted, "Readers may recall, if they use modems with their computers, that modem speeds are rated in 'bits per second', or bps. This is their channel capacity. Speed increases for modems plateaued at 56000 bps, because telephone connections are designed to have approximated that capacity, so no change in modem design can achieve higher speed." This indicates that physical constraints are very relevant in information processing and entropy theory of information from communication theory may play a larger role in understanding economic behaviors. Recently, a new theory of information was developed to expand Shannon's entropy theory of information into an economic theory (Chen, 2003, 2004, 2005). The basic idea of this theory is:

> Information is the reduction of entropy, not only in a mathematical sense, as in Shannon's theory, but also in a physical sense. The rules of information transmission developed in Shannon's theory, as mathematical rules, apply not only to communication systems, but also to all living organisms, including human beings.

By understanding information and the human mind from the foundation of physical laws, we are able to build up the theory of mind as a part of physics. Some earlier works have shown this has greatly simplified our understanding of mind (Chen, 2003, 2005, 2007). In the next two sections, we will present a mathematical theory of judgment as another example how intuition from statistical mechanics helps us gain deeper understanding of the mind.

II. Value of Judgment and Measure of Bias

The value of information is a function of probability and must satisfy the following properties:

(a) The information value of two events is higher than the value of each of them.
(b) If two events are independent, the information value of the two events will be the sum of the two.

The only mathematical functions that satisfy all the above properties are of the form

$$H(P) = k \ln P \quad (1)$$

where *H* is the value of information, *P* is the probability associated with a given event and *k* is a constant (Einstein, 1905; Applebaum, 1996). For simplicity, we will take *k* to be 1 in our paper. Formula (1) represents the level of uncertainty. When a signal is received, there is a reduction of uncertainty, which is information.

Suppose a random event, *X*, has *n* discrete states, $\{x_1, x_2, \ldots, x_n\}$, with probability $\{p_1, \ldots p_n\}$. The information value of *X* is the average of information value of each state, that is

$$H(X) = \sum_{j=1}^{n} p_j \ln(p_j) \quad (2)$$

The right hand side of (2), which is the entropy function first introduced by Boltzmann in the 1870s, is also the general formula for information (Shannon, 1948).

The subjective judgment of a person may differ from the objective probability. Suppose the subjective judgment of the probability distribution is $\{q_1, \ldots q_n\}$. Then the value of judgment on each q_i, is

(3)

The value of judgment of a random event is the average of value of judgment of each state, weighted by the objective probability distribution of the random event.

$$\sum_{j=1}^{n} p_j \ln(q_j) \quad (4)$$

Since the above expression will be used frequently, we will call it subjective entropy or subjective information.

In a random event with *n* possible outcomes, the state of equal distribution has the highest level of entropy. As the state of maximum entropy is the state of equilibrium, no further information input is required to determine the maximum entropy state. From the decision making perspective, the value of a judgment can be defined as the subjective entropy of that particular judgment minus the maximum entropy. That is

$$\begin{aligned} V(p,q) &= \sum_{j=1}^{n} p_j \ln q_j \quad - \sum_{j=1}^{n} \frac{1}{n} \ln(\frac{1}{n}) \\ &= \sum_{j=1}^{n} p_j \ln q_j \quad + \ln n \end{aligned} \quad (5)$$

When each

$$q_j = p_j, \qquad 1 \le j \le n$$

The value of judgment becomes the value of information.

$$\sum_{j=1}^{n} p_j \ln(p_j) \quad + \ln n \quad (6)$$

From Gibbs inequality (Gibbs, 1902; Isihara, 1971),

$$\sum_{j=1}^{n} p_j \ln(p_j) \ge \sum_{j=1}^{n} p_j \ln(q_j) \quad (7)$$

with equality achieved if and only if each

$$q_j = p_j, \qquad 1 \le j \le n$$

Therefore, the value of judgment is always less than or equal to the value of information with the same probability distribution. The difference between the value of information and the value of judgment is the measure of bias. Hence bias can be defined as

$$B(p,q) = \sum_{j=1}^{n} p_j \ln(p_j) - \sum_{j=1}^{n} p_j \ln(q_j) \quad (8)$$

which is always nonnegative and is zero if and only if each

$$q_j = p_j, \qquad 1 \le j \le n$$

In general, the difference between the first term and the second term of formula (8) is smaller when q_j is closer to p_j. Formula (8) is called relative entropy in information theory (Cover and Thomas, 2006).

For the simplicity of exposition, we will only consider events with two possible outcomes, state 1 and state 2 in the remaining part of the paper. When there are only two possible states, formula (5) and (8) become

$$V(p,q) = p\ln q + (1-p)\ln(1-q) + \ln 2 \qquad \textbf{(9)}$$

and

$$B(p,q) = p\ln p + (1-p)\ln(1-p) - \{p\ln q + (1\text{-p})\ln(1\text{-q})\} \qquad \textbf{(10)}$$

We will use some simple calculation to illustrate the properties of the value of judgment and the measure of bias. First, we will calculate the value of the neutral judgment, which will give 50% chance to both state 1 and state 2. From (9), the value of the neutral judgment is

$$p\ln 0.5 + (1-p)\ln 0.5 + \ln 2 = \ln 0.5 + \ln 2 = 0$$

Hence the value of neutral judgment is zero, regardless of the actual probability distributions of the states. Intuitively speaking, a neutral investor does not believe he possesses valuable information and puts his money into an index fund.

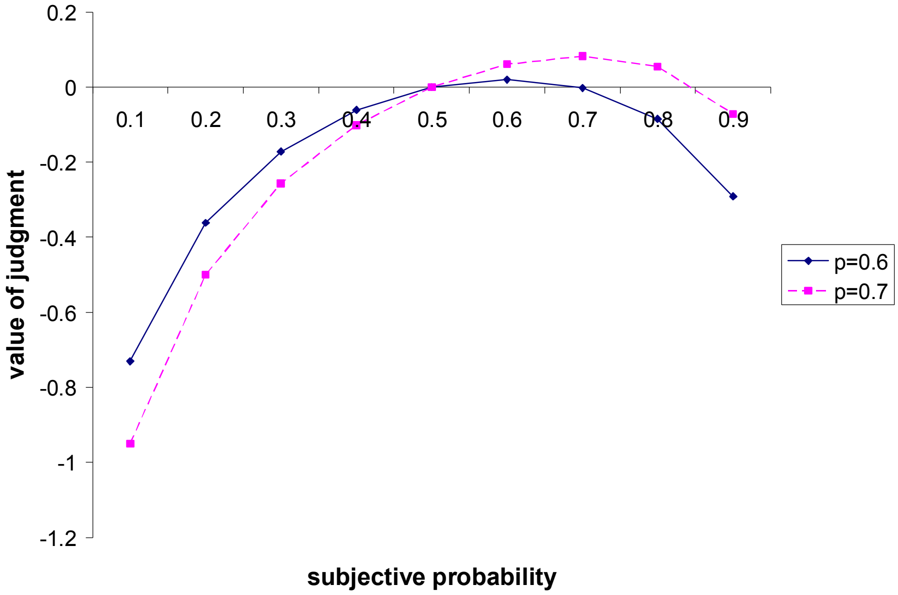

Figure 1. *Value of judgment*. The two curves are the values of judgment at different levels of subjective probability while objective probabilities are {0.6, 0.4} and {0.7, 0.3} respectively.

Figure 1 displays the values of judgment with different levels of subjective probability estimation when the objective probabilities are p=0.7, 1-p=0.3 and p=0.6, 1-p=0.4 respectively. From Figure 1, it can be seen that the value of a neutral judgment, i.e., q=0.5, is zero. When a judgment is at the wrong direction, that is, q<0.5 while p>0.5, the value of judgment is negative. When the judgment is precise, that is, q=p, the value of judgment reaches the maximum positive value. But even when a judgment is at the right direction, that is, q > 0.5 while p>0.5, the value of judgment could still be negative. Intuitively, suppose a risky asset will outperform the market moderately. If the asset is heavily weighted in a portfolio, the portfolio will still underperform the market on average. Comparing the value curves of p = 0.7 and p = 0.6, we find that when a random event is further away from equilibrium, that is, when p is further away from 0.5, there is more opportunity that the value of judgment will be positive and higher. This is consistent with the fact that a physical asset that is further away from equilibrium is more economically valuable.

Now consider two random events with different probability distributions. Assume in the first event, the probability of state 1 is 90% and the probability of state 2 is 10%. Someone estimates the probability of state 1 is 70% and the probability of state 2 is 30%. From (9), the value of his judgment is

$$0.9\ln 0.7 + 0.1\ln 0.3 + \ln 2 = 0.25$$

In the second event, the probability of state 1 is 70% and the probability of state 2 is 30%. Someone estimates the probability of state 1 is 70% and the probability of state 2 is 30%. From (9), the value of his judgment is

$$0.7\ln 0.7 + 0.3\ln 0.3 + \ln 2 = 0.08$$

From (10), the bias of judgment in the estimation of the first event is

$$0.9\ln 0.9 + 0.1\ln 0.1 - (0.9\ln 0.7 + 0.1\ln 0.3) = 0.12$$

while the bias of judgment in the estimation of the second event is zero.

From the above calculation, we find that, under certain conditions, the judgment that is more biased turned out to be more valuable than an unbiased judgment. Intuitively speaking, an investor who is modestly favorable to a stock which turns out to earn very a high rate of return will perform better than an investor who is modestly favorable to a stock which turns out to earn a moderate rate of return. In general, we have the following propositions.

Proposition 1

When q > 0.5, the value of judgment is an increasing function of p. When q < 0.5, the value of judgment is a decreasing function of p. Proof: Rewrite (9) into

$$V(p,q) = p(\ln q - \ln(1-q)) + \ln(1-q) + \ln 2 \quad (11)$$

This is a linear function of p. Note that

$\ln q - \ln(1-q) < 0$ when $q < 0.5$

and

$\ln q - \ln(1-q) > 0$ when $q > 0.5$.

The result becomes very clear.

Proposition 2

B(p, q) is an increasing function of p when $p > q$, and a decreasing function of p when $p < q$

Proof: Differentiate (10) with respect to p and simplify. We have

$$\frac{d}{dp}B(p,q) = \ln p - \ln q - (\ln(1-p) - \ln(1-q))$$

Figure 2 shows the value of judgment and measure of bias at different levels of p, assuming q is 0.6. From the figure, we can find that when the judgment is at the right direction, a judgment with a high level of bias is also of a high value.

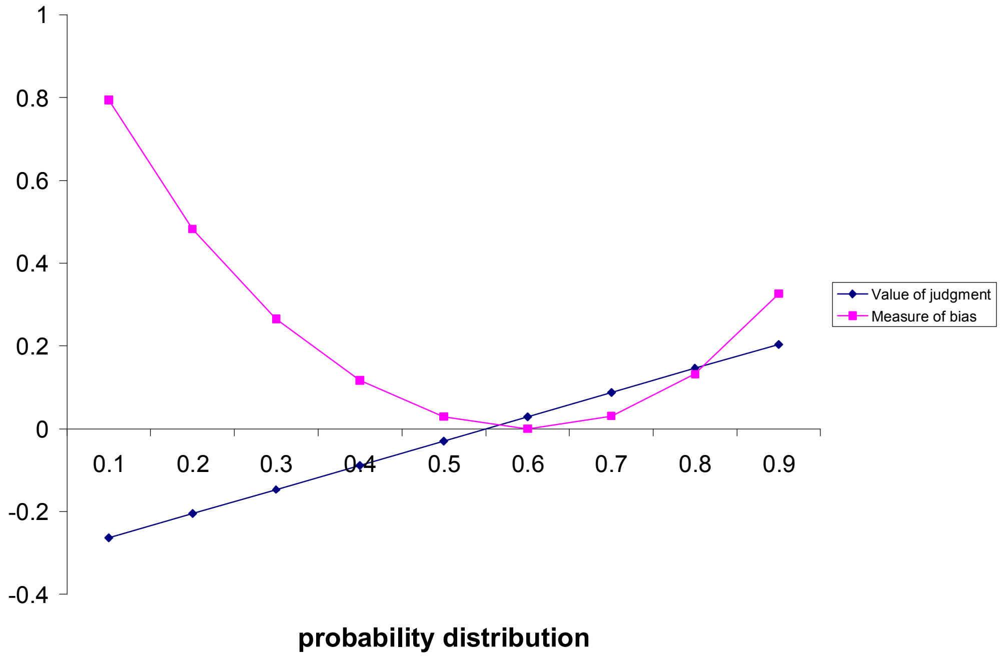

Figure 2. *Value of judgment and Measure of bias***.** The straight line represents the values of judgment and the curve is the measure of bias at different levels of objective probability. The subjective probability is {0.6, 0.4}. The measure of bias is zero when the objective probability equals the subjective probability.

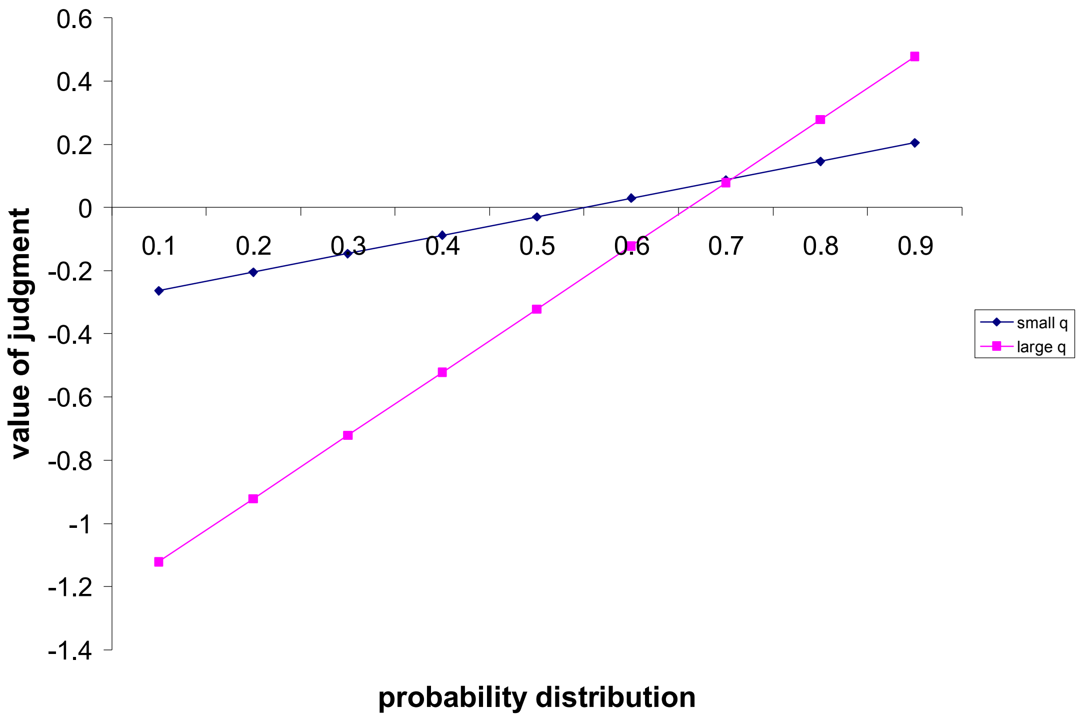

Figure 3. *Values of different judgments.* The two straight lines are values of judgment with different levels of subjective probability estimation. When subjective probability estimation deviates less from neutrality, the value of judgment turns positive earlier and the slope is gentler. When subjective probability estimation deviates more from neutrality, the value of judgment turns positive later and the slope is steeper.

We further investigate the structure of value distribution when people start making different kinds of judgments. Since the default judgment is $q = 0.5$, judgments can be classified by how far away q is from 0.5. Figure 3 draws two lines of value when q is equal to 0.6 and 0.8 respectively. When q is equal to 0.6, the value turns positive earlier and the slope is gentler. When q is equal to 0.8, the value turns positive later and the slope is steeper. So a more extreme judgment is riskier and when correct, has higher payoff. In general, we have the following,

Proposition 3
Assume $q > 0.5$. Similar results can be obtained for $q < 0.5$. V(0, q) is a decreasing function of q. V(1,q) is an increasing function of q. When q is smaller, the value of judgment turns positive earlier with respect to p and the slope is gentler. When q is larger, the value of judgment turns positive later with respect to p and the slope is steeper.
Proof: From (11),

$$V(0,q) = \ln(1-q) + \ln 2$$

and

$$V(1,q) = \ln q + \ln 2$$

Hence V(0,q) is a decreasing function of q and V(1,q) is an increasing function of q.
From (11), the slope of value of judgment as a function of p is

$$\ln q - \ln(1-q)$$

which is an increasing function of q.
From (11), the line of the value of judgment crosses x-axis at

$$p = \frac{\ln(1-q) + \ln 2}{\ln(1-q) - \ln q}$$

Differentiate p over q, we have

$$p'(q) = \frac{q \ln q + (1-q)\ln(1-q) + \ln 2}{q(1-q)(\ln(1-q) - \ln q)^2} > 0$$

Therefore, the value of judgment turns positive later when q is larger.

To further investigate the properties of the value of judgment, we will find the value of judgment averaged over p at each level of q. Integrating (11) over p from 0 to 1, we have

$$\int_0^1 (p(\ln q - \ln(1-q)) + \ln(1-q) + \ln 2\ )dp$$
$$= \frac{1}{2}(\ln q - \ln(1-q)) + \ln(1-q) + \ln 2 \qquad \textbf{(12)}$$
$$= \frac{1}{2}(\ln q + \ln(1-q)) + \ln 2$$

Figure 4 is the graph of the above result. The unconditional mean of value of judgment is always negative except at q = 0.5, at which point it is zero. This is a consequence of the entropy law, which states that entropy in a closed system moves towards its maximum. From the graph, the unconditional mean is smaller when q is further away from 0.5. This offers another explanation why people are in general conservative.

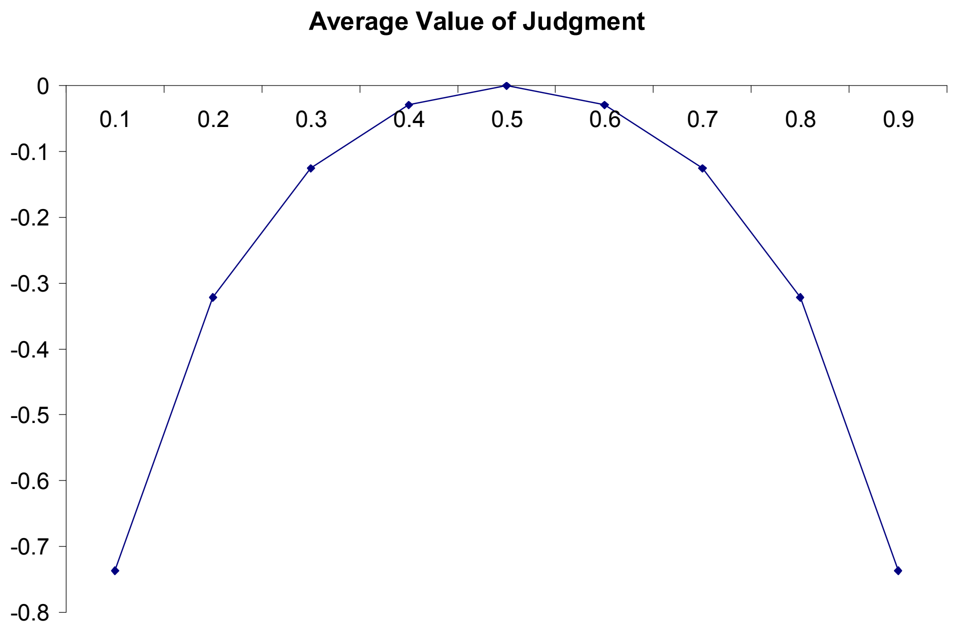

Figure 4. *Average values of judgment.* The curve is the value of judgment averaged over different levels of objective probability at each level of subjective probability. The unconditional mean of value of judgment is always negative except at q = 0.5, at which point it is zero. The unconditional mean is smaller when q is further away from 0.5. This offers another explanation why people are in general conservative.

III. Value of Judgment and the Expected Rate of Return of Portfolios

In this section, we will compare the value of judgment calculated from the theory with the expected rates of return of the portfolios constructed from the same judgment. Investment decisions are made according to investors' judgment about stocks. To quantify the relation between judgment and investment decision, we will consider a market with a risk free asset and some risky assets. A unit of the risk free asset provides a payoff of 1 at the end of each time period. This risk free asset is the default asset of a portfolio. Without loss of generality, all the risky assets provide payoff of 1 at period 0. In the next period, an asset's payoff may either increase by d or decrease by d, with different levels of probabilities. Assume that the market does not detect the differences among various risky assets, and their unit prices are identical and equal to the unit price of risk free asset.

We will assume investors hold simple portfolios with only two assets: a risk free asset and a risky asset. Investors aim at maximize expected geometric return (Fernholz, 2002; Sinn, 2003). Based on the subjective assessment of the return distribution of the risky asset, an investor can determine the optimal portion of the risky asset in the portfolio and calculate the expected rate of return of this portfolio. Suppose an investor believes one risky asset will have a probability q to increase its payoff by d and a probability 1- q to decrease its payoff by d. Assume the portfolio he constructed contains a portion x of risky asset and the remaining portion of $1 - x$ is risk free asset. The expected geometric return is

$$((1-x)+x(1+d))^q((1-x)+x(1-d))^{1-q}-1 \\ =(1+xd)^q(1-xd)^{1-q}-1 \quad (\mathbf{13})$$

To find out the value of *x* at which the portfolio will have the maximal rate of return, we differentiate the above formula with respect to *x*.

$$\frac{d}{dx}((1+xd)^q(1-xd)^{1-q}-1) \\ =d(1+xd)^{q-1}(1-xd)^{-q}(2q-1-xd)$$

The above differentiation equals zero when

$$x=\frac{2q-1}{d}$$

At this value of *x*, the portfolio obtains the highest expected geometric return. Plug the value of *x* into (13), the expected rate of return is

$$(1+2q-1)^q(1-(2q-1))^{1-q}-1 \\ =2q^q(1-q)^{1-q}-1$$

If the objective return distribution of the risky asset is {*p,* 1-*p*} instead of the subjectively assessed {*q,* 1-*q*}, the expected rate of return of the portfolio becomes

$$2q^p(1-q)^{1-p}-1 \quad (\mathbf{14})$$

The first order approximation of (14) is

$$\ln(2q^p(1-q)^{1-p}) \\ =p\ln q+(1-p)\ln(1-q)+\ln 2$$

Comparing the above result with (9), we find that the first order approximation of the expected rate of return of the portfolio constructed from a certain judgment is exactly equal to the value of the judgment defined in this work. This is not obvious that these two values should be equal. Initially, I was only looking for some numerical correlation between these two values.

Since the value of judgment provides a good approximation to the rate of return on investment, it can be conveniently used to understand the relation between human judgment and patterns in investment returns and stock market. Detailed discussion will be provides in a later work (Chen, 2008).

IV. Conclusion

This paper is a continuation of earlier works to understand human mind from statistical mechanics (Chen, 2003; 2005; 2007). We develop a mathematical theory on the value of judgment, which is generalized from the entropy theory of information. The elegant properties this theory exhibits are further indication that human mind is shaped by fundamental physical laws. An important property of this theory is that the value of a judgment provides a good approximation to investment return of the portfolio from the same judgment. The definiteness of this theory provides the discipline to the mostly ad hoc approach in behavioral models in investment.

References

Adami C. Information theory in molecular biology. Physics of Life Reviews 2004; 1:3-22

Applebaum D. Probability and Information, an integrated approach. Cambridge University Press, Cambridge, 1996.

Arrow K. Information and economic behavior, in Collected Papers of Kenneth J. Arrow, Volume 4, Cambridge, MA: The Belknap Press of Harvard University Press, (1973 (1983)).

Avery J. Information theory and evolution, World Scientific, Hackensack, NJ, 2003.

Baierlein R. Thermal physics, Cambridge University Press, Cambridge, 1999.

Bergstrom C and Lachmann M. Shannon information and biological fitness. Information Theory Workshop, IEEE, 2004; 50-54.

Chen J. An Entropy Theory of Psychology and its Implication to Behavioral Finance. Financiële Studievereniging Rotterdam Forum 2003; 6(1): 26- 31.

Chen J. Generalized Entropy Theory of Information and Market Patterns. Corporate Finance Review 2004; 9(3) :23- 32.

Chen J. The physical foundation of economics: An analytical thermodynamic theory. World Scientific, Hackensack, NJ, 2005.

Chen J. The Informational Theory of Investment: A Comparison with Behavioral Theories. ICFAI Journal of Behavioral Finance 2007; 4(1):6–31.

Chen J. Value of Judgment and Measure of Bias: A Common Framework for Behavioral and Informational Theories. Work in process, 2008.

Cover T and Joy T. Elements of information theory, 2nd edition, Wiley, Hoboken, NJ, 2006.

Einstein A. On a Heuristic Point of View Concerning the Production and Transformation of Light. In The Collected Papers of Albert Einstein, Volume 2, Princeton University Press, 1905 (1989).

Fernholz R. Stochastic portfolio theory, Springer, New York, 2002.

Gallager R. Claude E. Shannon: A Retrospective on his life, work, and impact, IEEE Trans. IT, 2001; 47: 2681-2695.

Gibbs JW. Elementary Principles in Statistical Mechanics, Ox Bow Press, Woodbridge, Connecticut, 1902 (1981).

Grossman S and Stiglitz J. On the impossibility of informationally efficient markets. American Economic Review 1980; 70:393-408.

Isihara A. Statistical physics, New York: Academic Press, 1971.

Jaynes ET. Information Theory and Statistical Mechanics. Physical Review 1957; 106: 620-630

Maxwell J. Theory of heat. Green & Co., Longmans, London, 1871.

Peng L and Wei X. Investor attention, overconfidence and category learning. Journal of Financial Economics 2006; 80:563-602.

Peng L. Learning with information capacity constraints. Journal of Financial and Quantitative Analysis 2005; 40:307-330.

Qian H. Relative entropy: Free energy associated with equilibrium fluctuations and non-equilibrium deviations. Physical Review E 2001; 63:042103.

Shannon C and Weaver W. The mathematical theory of communication. Urbana: The University of Illinois Press, 1949.

Shannon C. A mathematical theory of communication. The Bell System Technical Journal 1948; 27: 379-423, 623-656.

Shannon C. The Bandwagon (Editorial), IEEE Transactions Information Theory 1956; 2:3.

Sims C. Implication of rational inattention. Journal of Monetary Economics 2003; 50: 665-690.

Sinn HW. Weber's law and the biological evolution of risk preferences: The selective dominance of the logarithmic utility function. The Geneva Papers on Risk and Insurance Theory 2003; 28:87-100.