An Entropy Theory of Psychology and its Implication to Behavioral Finance

Jing Chen
School of Business
University of Northern British Columbia
Prince George, BC
Canada V2N 4Z9
Phone: 1-250-960-6480
Fax: 1-250-960-5544
Email: chenj@unbc.ca
Web: http://web.unbc.ca/~chenj/

Abstract

Most of the theories in behavioral finance rely on some kind of psychological biases. However, the potentially boundless set of psychological biases that theorists can use to build behavioral models and explain observed phenomena creates the potential for "theory dredging." We develop a unified theory of human psychology based on entropy law, the most universal natural law. This unified theory of human psychology will help us determine whether patterns discussed in the finance literature are genuine or the result of data mining. It will also greatly reduce the possibility of "theory dredging" in the future works on behavioral finance.

An Entropy Theory of Psychology and its Implication to Behavioral Finance

Because of the inability of efficient market theory to explain the persistence of some patterns in the financial markets, many new theories have emerged to understand these patterns. Most of these theories rely on some kind of human psychological biases and are generally grouped under the category of behavioral finance. However, "the potentially boundless set of psychological biases that theorists can use to build behavioral models and explain observed phenomena creates the potential for 'theory dredging.'" (Chan, Frankel and Kothari, 2002) Thus it is difficult to distinguish data mining from genuine patterns. It would be very helpful to develop a unified theory of human psychology based on a sound foundation to understand market patterns.

The patterns in financial markets reflect the patterns of information processing by the investment public. Since information is the reduction of entropy and all human activities are essentially entropy processes, it is natural to understand human psychology and market patterns from the viewpoint of entropy theory.

Recently, Chen (2003) showed how entropy theory provides clear understanding of some of the most common patterns in psychology and financial markets. In this paper, we will provide a systematic discussion about the patterns of human psychology and explain how entropy theory offers a simple and unified understanding of these patterns. Since entropy process is the most universal physical process, a unified framework built on entropy theory will greatly reduce the possibility of "theory dredging" in the future works on behavioral finance. It will also help us determine whether some patterns frequently discussed in the finance literature, such as equity premium puzzle, are genuine or the result of data mining.

The remainder of the paper is structured as follows. Section I introduces the entropy theory of information. Section II explains how entropy theory offers a unified understanding of the patterns of human psychology. In section III, we apply the entropy theory of psychology to examine whether a perceived pattern in the financial market is genuine or the result of data mining. Section IV concludes.

I. The entropy theory of information

The concept of information is intimately related to entropy. In a thought experiment, Maxwell (1871) reasoned, if information is costless, the entropy of a system can be decreased, which violates the second law of thermodynamics. He concluded that the physical cost of obtaining information must be at least as much as the value of information. In 1870s, Boltzmann defined the mathematical function of entropy, which Shannon (1948) identified as information many years later. The works by Shannon and others explicitly established the equivalence between information and physical entropy. (Bennett,1988). In the following, we give a brief introduction of information theory.

The value of information is a function of probability. It satisfies the following properties:

1. The information value of two events is higher than the value of each of them.

2. If two events are independent, the information value of the two events will be the sum of the two.

3. The information value of any event is non-negative.

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

$$H(P) = -\log_b P \qquad (1)$$

where b is a positive constant. (Applebaum, 1996) 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, …,x_n, each with probability p_1, p_2, …,p_n. The information value of X is the average of information value of each state, that is

$$H(X) = -\sum_{j=1}^{n} p_j \log(p_j) \qquad (2)$$

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

Figure 1 is a graph of formula (1), where H is a function of P, the probability of event. From Figure 1, value is a decreasing function of probability. When P = 1, -log P =0. The value of information that is known to everyone is zero. When P approaches zero, -log$P$ approaches infinity. The value of information that is known to few is very high. For example, if an unexpected surge of corporate profit is known to very few people privately, i.e., when P is very small and –log P is very big, the information is highly valuable. Huge profit can be made by trading the underlying stocks. But if the information is announced publicly and is known to many people, the value of information is very low. Little profit can be made from trading on such information.

It is often said that the cost of information has dropped sharply over the years. But at the same time, the value of the same type of information has dropped sharply as well. Because of the high value of information that is only known or understood by few, important information is carefully guarded. For example, Warren Buffett will not announce to the public which stock he is going to buy shortly.

Even if information is distributed freely, a receiver may not be able to comprehend its full meaning. Following Shannon (1948), the amount of information one can receive, R, would be equal to the amount of information sent minus the average rate of conditional entropy.

$$R = H(x) - H_y(x) \tag{3}$$

The conditional entropy $H_y(x)$ is called the equivocation. It measures the average ambiguity of the received signal. So $H_y(x)$ is the quantitative measure of information asymmetry. When x and y are independent, $H_y(x) = H(x)$ and $R = 0$. No information can be transmitted between two objects who are independent to each other. This shows that it is very difficult for most people to understand the value of a new idea, product or organizational structure. In general, the amount of information one can receive from the source depends on the correlation between the two. The higher the correlation between the source and receiver, the more information can be transmitted. Since different people have different background knowledge about the same information, heterogeneity of opinion occurs naturally.

The entropy theory of information provides a quantitative measure of information asymmetry and shows that the cost of obtaining some information is positively correlated with the value of information. This greatly refines the current information theory in economics, where information asymmetry is a qualitative concept and the cost of obtaining information is generally determined exogenously. (Grossman and Stiglitz, 1980)

II. Human psychology: An entropy perspective

The entropy law, which states that closed systems tend towards a maximum entropy equilibrium state, is the most universal law of nature. From entropy law, it is far easier for a system to disintegrate than to maintain its structure. (Morowitz, 1992; Margulis, 1998) There is, therefore, a strong selective pressure for important knowledge to become genetically coded into heuristic principles to reduce the cost of learning. (Tversky and Kahneman, 1974) For human beings, these heuristic principles are called human psychology. We will discuss some of the frequently cited patterns of human psychology in behavioral finance literature and offer a unified understanding from the entropy perspective.

1. Conservatism

Conservatism states that individuals update their beliefs slowly in the face of new information. This property is a natural result from information theory. From (3), the information one can receive is information sent minus equivocation, which is reduced gradually as the receiver’s background knowledge about the source increases. Hence conservatism reflects the gradual reduction of equivocation by the receiver of information.

2. Framing or representativeness

We often frame problems into categories and assign different categories based on different level of importance. Why do we do that? The following result from statistical physics helps answer this question.

If $\{p_1, \ldots p_n\}$ and $\{q_1, \ldots q_n\}$ are two sets of probabilities, then

$$-\sum_{j=1}^{n} p_j \log(p_j) \leq -\sum_{j=1}^{n} q_j \log(p_j) \qquad (4)$$

with equality if and only if each

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

This result is called Gibbs inequality. (Isihara, 1971) In Gibbs inequality, p_j can be understood as the probability of event j in nature and q_j is the subjective probability of our assessment of that event. The left hand side of formula (4) is the average uncertainty of events and the right hand side the uncertainty of our assessment of those events. In general, the difference between the left hand side and right hand side of (4) is smaller when q_j is closer to p_j. This means information processing is more efficient when the subjective probabilities are closer to the objective probabilities. In particular, a mind with stored data about the natural environment is in general more efficient than a completely unbiased mind, where all subjective probabilities are to be learned from scratch. Natural selection determines that the human mind will evolve so that, “in general, instances of large classes are recalled better and faster than instances of less frequent classes; that likely occurrences are easier to imagine than unlikely ones; and that the associative connections between events are strengthened when the events frequently co-occur.” (Tversky and Kahneman, 1974, p.1128)

“People rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations. In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors.” (Tversky and Kahneman, 1974, p.1124) What causes these severe and systematic errors? Human minds are the result of natural selection, which “operates over thousands of generations. For ninety-nine percent of human existence, people lived as foragers in small nomadic bands. Our brains are adapted to that long-vanished way of life, not to brand-new agriculture and industrial civilizations.” (Pinker, 1997, p. 42) This is why we observe systematic errors in judgment of human beings from today’s perspective. For example, most of us still have a great fear of snakes, although they rarely pose a threat to urban dwellers today. On the other hand, the fear of electricity has to be instilled into children’s minds with great difficulty. (Pinker, 1997)

3. Herd behavior

From the second law of thermodynamics, a random action generally costs more than it gains. To concentrate actions into profitable ones, we often learn from the experience of successful people and copy their behavior. It is generally very costly or impossible to repeat all experiences of others to investigate the reasons behind a certain action. Copying the actions of others directly is much easier. Herding mentality developed because it is a cost-effective way of learning most of the time.

4. Overconfidence

"Extensive evidence shows that people are overconfident in their judgments." (Barberis and Thaler, 2003) From entropy law, any biological system, as a non-equilibrium system, faces constant dissipation. Endless efforts are required to maintain a non-equilibrium system. Entropy law has been intuitively understood since ancient times. "The gods had condemned Sisyphus to ceaselessly rolling a rock to the top of a mountain, whence the stone would fall back of its own weight. They had thought with some reason that there is no more dreadful punishment than futile and hopeless labour. … If this myth is tragic, that is because its hero is conscious. … The workman of today works every day in his life at the same tasks and this fate is no less absurd. But it is tragic only at the rare moments when it becomes conscious." (Camus, 1955, p. 109) In the course of the evolution of the solar system, all life on earth will go extinct in the distant future. (Lovelock, 1988) From a purely rational perspective, life is meaningless. Since human beings are self-conscious, why life is worth living always lingers in people's minds. "There is but one truly serious philosophical problem and that is suicide. Judging whether life is or is not worth living amounts to answering the fundamental question of philosophy." (Camus, 1955, p. 11) Overconfidence is an adaptive psychological trait that helps us survive in this world.

5. Loss aversion in winning and risk seeking in losses

Human beings often exhibit loss aversion in winning and risk seeking in losses. Kahneman and Tversky (1979) collected some responses to hypothetical choice problems. In one problem, the subjects were presented with two choices.

Choice A: There is an 80% probability of winning 4000 pounds and a 20% probability of winning nothing.

Choice B: There is a certainty of winning 3000 pounds.

The expected end wealth of choice A is 3200 and of choice B is 3000. Most respondents chose B, exhibiting loss aversion in winning. When the signs of the outcomes are reversed, the problems become the following:

Choice C: There is an 80% probability of losing 4000 pounds and 20% probability of losing nothing.

Choice D: There is a certainty of losing 3000 pounds.

The expected end wealth of choice C is -3200 and of choice D is -3000. Most respondents chose C, exhibiting risk seeking in losses. As money is a new invention in human evolutionary history, the preference for money must be derived from something else. Since food is the most important resource of our evolutionary past, our preference for wealth is probably derived from our preferences for food.

In the history of human evolution, we have not been able to store large amounts of extra food. If one goes without food for several days, he will starve. We translate the monetary numbers from the above four questions into days of food to obtain the following. In the case of gain, we can think of the choices of two possible strategies. In the first strategy, there is an 80% probability of getting food for 40 days and 20% chance of getting nothing. In the second strategy, there is a certainty of getting food for 30 days. It is easy to see why most people will prefer 30 days of food in certainty over a strategy that contains a 20% risk of getting nothing. In the case of loss, we can think of the choices of two possible strategies. In the first strategy, there is an 80% probability of getting no food for 40 days and a 20% chance of getting nothing for zero day. In the second strategy, there is a certainty of getting no food for 30 days. Since 30 days' without food will represent sure death, people will naturally choose 20% chance of survival. So people consistently avoid risk in both positive gain and negative loss. "Risk seeking" in loss is an unfortunate terminology borrowed from utility theory.

From the above discussion, we find that some psychological patterns, such as conservatism, reflect the constraint of physical laws. Others, such as framing and herding, are evolutionary adaptations to enable the efficient processing of information. Still others, such as overconfidence and loss aversion, are mental attitudes that help us survive the constant dissipation endured by all non-equilibrium systems. Therefore, the entropy theory offers a unified understanding of human psychology.

III. **Market patterns or data mining**

A unified theory of human psychology based on the foundation of fundamental physical laws will help us distinguish genuine patterns in the financial markets from data mining. In this section, we will discuss the equity premium puzzle.

Mehra and Prescott (1985) observed that the large size of risk premium on US equities can not be explained by the standard general equilibrium models and called it a puzzle. Among much research generated by this observation, two approaches are relevant to our study. One approach attributes the high risk premium to loss aversion by investors. Barberis and Thaler (2003) provided a survey of works along this line. The other is survivorship bias proposed by Jorion and Goetzmann (1999).

From the entropy theory of information and human psychology, how long a pattern persists depends on the cost of learning. It is often very costly to gain a deep understanding of a company or an industry, especially when the industry is new. The equity premium puzzle, however, is a very simple pattern on financial data, which, once discovered, can be understood very easily by the investment public. The strategy of profiting from the high equity premium is easy to implement and of low risk. This indicates that the pattern of high equity premium, if it does exist, is a short term one.

Whether high equity premium is a pattern or a result of selection bias can be answered by more comprehensive data. Jorion and Goetzmann (1999) documented that among all the equity markets around the world in the past century, the US market had the highest return. They argued that US market had the highest return because US was the most successful economic system in the world in the last century. “For 1921 to 1996, U.S. equities had the highest real return of all countries, at 4.3 percent, versus a median of 0.8 percent for other countries. The high equity premium obtained for U.S. equities appears to be the exception rather than the rule.” (Jorion and Goetzmann,1999, p. 953) Their conclusion is consistent with the entropy theory of information and human psychology.

IV. Concluding remarks

Very often, it is in a data driven and highly technical subject that revolutionary ideas originate. Modern astronomy was a data driven and highly technical subject aimed to understand the movements of several planets. Newtonian mechanics, which was originally developed to provide a physical foundation of celestial movements, has become to dominate the thinking of social sciences for many years. Life processes, however, are thermodynamic processes instead of mechanical processes. Social processes, as life processes of one species, should be built on the theory of thermodynamics instead of Newtonian mechanics.

Finance is a very distinct subject in social sciences. The vast amount of reliable data accumulated in the financial market enable us to test different hypotheses at much higher confidence levels than in other subjects in the social sciences. At the same time, the wealth of information extracted from the financial data forces us to rethink the foundations of many theories that are related to financial economics. This work shows how problems in finance, a data driven and highly technical subject, stimulate researches in much broader areas.

Bibliography

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

Barberis, N. and R. Thaler, 2003, A survey of behavioral finance, in *Handbook of the economics of finance*, edited by G. Constantinides, M. Harris and R. Stulz.

Bennett, C. H., 1988. Notes on the History of Reversible Computation. IBM Journal of Research and Development, 32, 16-23.

Camus, A., 1955, The Myth of Sysiphus, Penguin Books, London.

Chan, W., Frankel, R. and Kothari, S. 2002, Testing Behavioral Finance Theories Using Trends and Sequences in Financial Performance, Working paper.

Chen, J. 2003, The Physical Foundation of Human Psychology and Behavioral Finance, Working paper.

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

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

Jorion, P. and Goetzmann, W. 1999, Global Stock Markets in the Twentieth Century, Journal of Finance, 54, 953-980.

Kahneman, D. and A. Tversky, 1979, Prospect theory: A analysis of decision under risk, Econometrica 47, 263-291.

Margulis, L., 1998. Symbiotic Planet, A new view of evolution, Basic Books, New York.

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

Mehra, R. and Prescott, E. 1985, The Equity Premium: A puzzle? Journal of Monetary Economics, 22, 133-136.

Morowitz, H., 1992, Beginnings of cellular life: metabolism recapitulates biogenesis, New Haven: Yale University Press

Pinker, S., 1997, *How the mind works*, (W. W. Norton. New York).

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

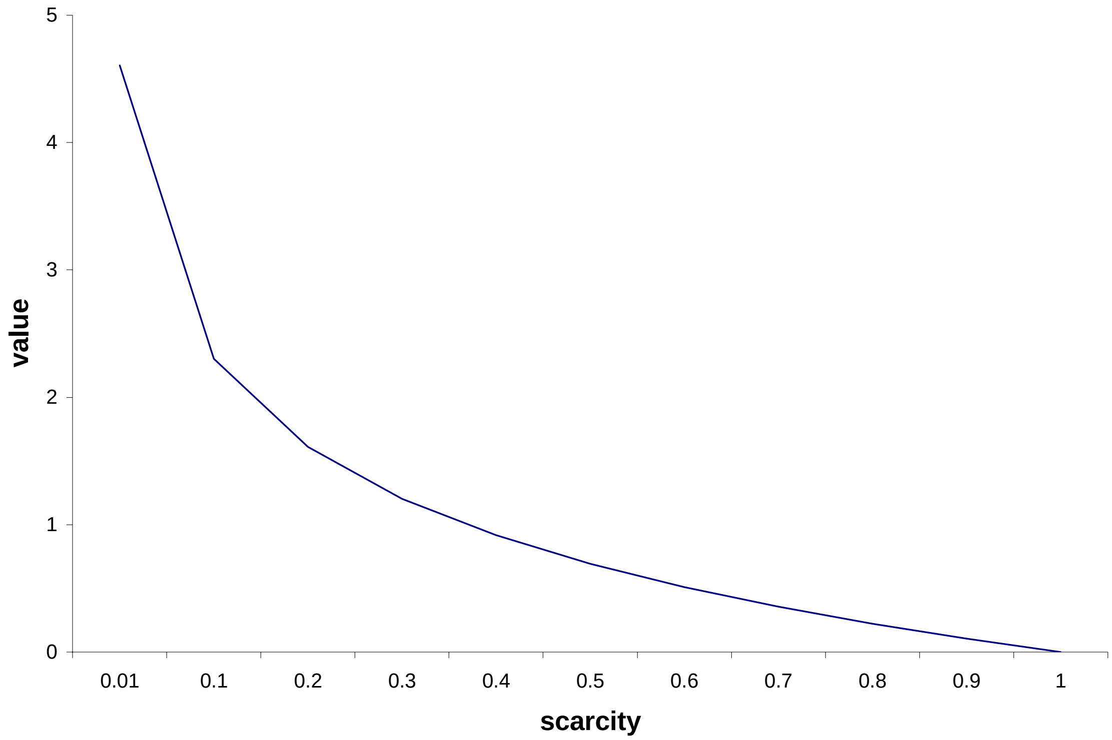
Prigogine, I., 1980. From being to becoming : time and complexity in the physical sciences. W. H. Freeman, San Francisco

Schrodinger, E., 1944. What is Life? Cambridge University Press, Cambridge

Tversky, A. and D. Kahneman, 1974, Judgment under uncertainty: Heuristics and biases, Science 185, 1124-1131.

Figure Captions

Figure 1: Value and scarcity



5
4
3
2
1
0
value
0.01
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
scarcity