AN INFORMATION THEORY OF ASSET PRICE DISTRIBUTIONS

By

J. Orlin Grabbe

Working Paper No. 3-83

RODNEY L. WHITE CENTER
FOR FINANCIAL RESEARCH

The Wharton School
University of Pennsylvania
Philadelphia, PA 19104

An Information Theory of Asset Price Distributions

by

J. Orlin Grabbe

The Wharton School
The University of Pennsylvania

July 1982
Revised: January 1983

An Information Theory of Asset Price Distributions

J. Orlin Grabbe

ABSTRACT

The Theil-Barbosa framework of "rational random behavior" in the theory of consumer demand involves a decision-maker who minimizes the sum of a loss functional u and the cost of information c(I). The definition of information I is that used in the information theory literature associated with Shannon, Wiener, Kullback, and Leibler. The present paper shows such a framework is appropriate for modelling the behavior of traders in speculative markets who announce bid-ask prices at which they are willing to trade. Traders will announce prices which will be drawn from a probability distribution, the selection of which is a function of the loss functional u and the cost of information. The lognormal distribution is derived as a special case. It is shown that the variance of such distributions depends on the marginal cost of information, but that the kurtosis depends on the shape of the loss functional. When the loss functional is less convex than a quadratic, distributions will be leptokurtic.

An Information Theory of Asset Price Distributions

J. Orlin Grabbe

I. Introduction

Why do prices in asset markets have one probability distribution as opposed to another? The question of what price to announce is the central problem faced by a trader in speculative markets like those for foreign exchange or eurocurrency deposits. Presumably price announcements are based on rational criteria, and these criteria will determine what distribution a time series of such prices will exhibit ex post. Why do price distributions sometimes have a lognormal distribution and sometimes not? Under what circumstances would we see an exponential or a dirichlet distribution?

The answer given here is that a trader preselects the probability distribution from which he will draw prices. The framework assumed is the same as used by Barbosa (1975) and Theil (1980) in connection with consumer choice. Theil derives the normal distribution from a quadratic loss functional, and Barbosa derives the dirichlet distribution from a different loss functional. Crucial to these results is the notion of minimum information (maximum entropy) as derived from information and communication theory. It should be emphasized at the outset that by "information theory" is meant the tradition stemming from Shannon (1948), Wiener (1962), and Kullback and Leibler (1951), and not the more recent economic tradition, although there are connections between the two.

In this paper I solve the trader's problem in general form for all probability distributions with a continuous density. The solution is applied to some specific cases. The role of "information measure" is clarified, and the lognormal is derived as a minimum information distribution. Though it is well known in the information theory literature that the normal distribution occurs as a minimum information distribution (Dowson and Wragg, 1973, for example), it is apparently not known that the same is true for the more financially-relevant lognormal distribution. Finally, circumstances in which asset returns will exhibit leptokurtosis are pointed out. One of these may provide an explanation for the leptokurtosis typically found in such returns in financial markets.

It seems appropriate to begin with an extended example. Consider the traditional story of the Walrasian auctioneer who announces prices in an auction setting. Supply and demand, based on the announced price, are tabulated. If these are not equal, a new price is announced, and the process continues until an equilibrium is reached. Trading then takes place at the equilibrium price. A real-world example of a Walrasian auctioneer is provided by the twice daily price-fixing in the London gold market. Representatives from five major trading houses meet at 10:30 A.M. and 3:00 P.M. in the fixing room at N. M. Rothschild. Each representative is in contact with his firm's trading room, which has orders collected from its customers all over the world to buy or sell gold at various prices. An auctioneer calls out a sequence of prices per ounce of gold, and each of the representatives tabulates the net amount his firm is willing to either buy or sell at that price. The price is fixed when excess demand is zero. Usually the fixing price is established fairly quickly, but on one occasion in October 1979 it took one hour and thirty-nine minutes to reach an equilibrium across the five trading houses.

Real-world markets, however, such as those for foreign exchange or eurocurrency deposits, typically differ from the Walrasian framework in at least three respects. First, the auctioneer, or trader, announces not one price, but two: a <u>bid</u> price at which he is willing to buy and an <u>asked</u> price at which he is willing to sell. Second, there is not one trader, but hundreds, all of which announce prices simultaneously. Thirdly, none of the traders knows the equilibrium market price prior to trading. Rather, each trader has to infer the nature of excess demand based on the frequency at which he is "hit" at the asked price relative to the frequency at which he is hit at the bid price. Equilibrium is a local phenomenon from a trader's perspective.

Suppose that, within this setting, we impose the following assumptions. At any point in time, there is a price $\overline{x}$ such that if a trader calls out bid and ask prices, x_b, x_a, at equal fixed increments δ from $\overline{x}$, then such price announcements will be optimal according to some objective function. To simplify the discussion, explicit reference to $x_a = x + \delta$ and $x_b = x - \delta$ will be dropped, and we will say "the trader announces price x," where x is the middle price between x_a, x_b. Therefore, if the trader announces price $x = \overline{x}$, then he is at an optimum, while if he announces $x \neq \overline{x}$, there is a loss incurred by departure from this optimum. Our first assumption, then, is the existence of such a loss functional u:

$$u = u(x, \overline{x}) .$$

Secondly, assume that the price x the trader announces will be drawn from some probability distribution F(x). Our objective is to determine what probability distribution the trader will use to generate price announcements, given his loss functional $u(x, \overline{x})$. Our problem is not, however, sufficiently well structured at this point to allow us to answer this question. We have to introduce the notion of information. The sequence of prices the trader calls out will be based on information about the location of the optimal price $\overline{x}$. What, precisely, do we mean by information in this context? Let us make a brief digression into information theory.

The definition of information occurs in the literature on information theory in association with discrete and continuous probabilities. Let X be a random variable that takes on values x_i with probabilities p_i, $i = 1, \ldots, n$. Then the information of X, S(X), defined as the negative of the discrete entropy of Shannon (1948), is

$$S(X) = \sum_{i=1}^{n} p_i \ln p_i \;.$$

This has a clear interpretation as a measure of the sharpness of the distribution, since S(X) takes a maximum value of 0 when one p_i is equal to one and all others are zero,[1] and a minimum value of $-\ln n$ when all probabilities are equal.

Thus, if we were to apply this measure of information to our trader, we would say that "the trader calls out price x_i with the probability p_i." Then the case S(X) = 0 would correspond to the case the trader called out a single one of the prices with certainty. The trader would only do so if he had maximal knowledge that he was calling out the optimal price. By contrast, the case $S(X) = -\ln n$ would correspond to the case that the trader would call out any one of the prices with equal probability. This would mean that the trader had no knowledge to indicate that one price was better than another.

The information S is defined for random variables with a discrete distribution. But we may extend the Shannon definition in such a way that it applies to arbitrary distributions. The mathematical procedure for doing so is described in Everett (1973). One such definition for a continuous random variable Y, suggested by Shannon as well as Wiener (1961), is

$$W(Y) = \int_r f(y) \ln f(y) dy$$

where r is the sample space. When r is a finite interval [a, b] of the real line, the distribution having the smallest amount of information is simply the uniform distribution. Again we note that minimal information in this context would imply the trader has no knowledge that suggests a price selected from one point in the interval is superior to any other point. By contrast, greater knowledge would enable the trader to concentrate his price drawings in a segment of the interval.

An axiomatic characterization of W(Y) may be found in Campbell (1972). A principal difference between the discrete information S and the continuous information W, as noted in Kolmogorov (1956) and Everett (1973), is that W is not invariant to the units used to measure y. That is, W is not invariant to coordinate transformation in the space of r. This suggests we should use an "information measure" relevant to the economic problem at hand, and leads to the following:

<u>Definition</u>: Let X be a random variable over a subset r of the real line with a continuous density f(x), and let g(x) be a non-negative function ("information measure") such that $g(y) = 0$, $y \varepsilon r$, implies $f(y) = 0$. Then the information $I_g(X)$ (relative to the information measure g(x)) is

$$I_g(X) = \int_r f(x) \ln \frac{f(x)}{g(x)} dx \; . \tag{1}$$

In particular, we will use the information measures $g(x) = 1$, $g(x) = \frac{1}{x}$, giving rise to two corresponding definitions of information:

$$I(X) = \int_r f(x) \ln f(x) dx \tag{2A}$$

$$I^*(X) = \int_r f(x) \ln\big(x f(x)\big) dx \; . \tag{2B}$$

The idea of using alternative information measures here was inspired by Ingarden and Kossakowski (1971), who showed that the Poisson distribution could be derived as a minimum information distribution when a variate, S*(X), of the discrete information of Shannon is used as the definition, namely

$$S^*(X) = \sum_{i=0}^{\infty} p_i \ln(i!p_i) .$$

Similarly, by using the definition I*(X), I am able in this paper to derive a class of minimum-information distributions that includes the lognormal as a member. (The "rule" governing the creation of I* may be ascertained from Lemma 3, below.)

Now we return to the problem faced by our trader. The trader does not know what the optimal price $\overline{x}$ is. But he can get an idea. He can observe the relative frequency at which he trades at the bid price versus the asked price. He can do research on the state of "fundamental" variables that affect supply and demand. The amount of information he has will be reflected in his choice of distribution F(x). I will assume throughout the paper that F(x) has a continuous density f(x) dx. Then information may be measured by I_g, and the cost of obtaining this information by $c(I_g)$. I will also assume that $c(I_g)$ is differentiable.

Thus the trader's problem is to choose a probability distribution from which to draw prices. His choice of a density f(x) dx will determine the expected loss

$$\overline{u} = \int_r u(x, \overline{x})\, f(x)\, dx$$

and the cost of information

$$c(I_g) = c\left(\int_r f(x) \ln \frac{f(x)}{g(x)}\, dx \right) .$$

If we measure $\overline{u}$ and $c(I_g)$ in the same units, then the trader's problem may be represented as

$$\min_{f(x)} \; \overline{u} + c(I_g) \; . \tag{T}$$

For the remainder of the paper, any probability distribution whose density function is a solution to (T), for given u and I_g, will be referred to as a "minimum information distribution." In addition, all prices x will be measured either as deviations from the hypothesized optimal value $\overline{x}$, or else as proportions of $\overline{x}$.

Obviously the loss function u can take many forms. For example, if the trader has a utility function which is maximized at $\overline{x}$, then the loss function u may be defined as the excess of the maximum utility level over the level actually obtained. The forms of u(x) investigated in this paper are summarized at the beginning of section III.

II. Some Preliminary Lemmas on Minimum Information

We now proceed to some preliminary lemmas that form the basis of the theorems in Section III. Lemma 2 gives a fairly general solution for the density function of minimum-information distributions with continuous densities. Lemma 1 is then used to prove that if a minimum information distribution exists, it is unique. (If u(x) is bounded, it is easy to prove that a minimum information distribution always exists. However, the functional forms for u(x) that I use later are all unbounded, so I ignore this special case.) Lemma 3 then demonstrates a correspondence among transformations of u(x), $I_g(X)$, and the minimum-information density.

Lemma 1: Let r be a subset of the real line, and let u be a nonnegative function defined on r, such that the integral

$$G(b) = \int_r e^{-bu(w)}dw$$

is finite for any positive b. Then if the equation

$$\frac{G'(b)}{G(b)} = -\overline{u} \tag{3}$$

has a solution, that solution is unique.

<u>Proof:</u> Rewriting equation 3, we have

$$0 = G'(b) + \overline{u}G(b) = \int_r \left(\overline{u}-u(w)\right)e^{-bu(w)}dw \; .$$

Multiplying by $e^{b\overline{u}}$, this becomes

$$0 = \int_r \left(\overline{u}-u(w)\right)e^{b(\overline{u}-u(w))}dw = \int_r v(w)e^{bv(w)}dw$$

$$= F(b), \text{ where } v(w) = \overline{u}-u(w) \; .$$

Now $F'(b) = \int_r v^2 e^{bv}dw > 0$, so that F is strictly increasing in b. Hence if $F(b) = 0$ has a solution, it is a unique solution. But this means $G'(b) + \overline{u}G(b) = 0$ has a unique solution. #

<u>Lemma 2:</u> Let u(x) be a non-negative function of a continuous random variable X, defined on a subset r of the real line. Then the probability density function f(x) which minimizes $c(I_g) + \overline{u}$, or

$$c\left(\int_r f(x)\ln\frac{f(x)}{g(x)}\,dx\right) + \int_r u(x)f(x)dx \tag{4}$$

is

$$f(x) = \frac{g(x)}{G(b)}\,e^{-bu(x)}$$

where $G(b) = \int_r g(x)e^{-bu(x)}dx$, and where $b = \frac{1}{c'}$ is the unique solution to $\frac{G'(b)}{G(b)} = -\overline{u}$, provided such a b exists and G(b) is defined.

Proof: We may approximate c locally at any point by $c = c_0 + c'\int_r f(x)\ln \frac{f(x)}{g(x)}\,dx$. We want to maximize $-c - \overline{u}$ subject to

$$\int_r f(x)dx = 1 \; . \tag{5}$$

Let λ be a Lagrange multiplier. We will maximize the Lagrangian using the well-known inequality that

$$\ln x < x-1 \quad \text{if} \quad x \neq 1$$
$$\ln x = x-1 \quad \text{if} \quad x = 1$$

From (4), (5), the Lagrangian L becomes, for $\lambda_1 = \lambda/c'$,

$$L - c_0 = c'\int_r f(x)\left(\ln \frac{g(x)}{f(x)} - \lambda_1 - \frac{1}{c'}\,u(x)\right)dx$$

$$= c'\int_r f(x)\ln\left(\frac{g(x)}{f(x)}\,e^{-\lambda_1 - \frac{1}{c'}u(x)}\right)dx$$

$$\leq c'\int_r f(x)\left(\frac{g(x)}{f(x)}\,e^{-\lambda_1 - \frac{1}{c'}u(x)} - 1\right)dx \; .$$

The equality holds if and only if $f(x) = g(x)e^{-\lambda_1 - \frac{1}{c'}u(x)}$. From (5), we then have

$$1 = e^{-\lambda_1}\int_r g(x)e^{-\frac{1}{c'}u(x)}dx = e^{-\lambda_1}G(\tfrac{1}{c'}) \; ,$$

provided the integral $G(\frac{1}{c'})$ is defined. If so, this implies

$$\lambda_1 = \ln \, G(\tfrac{1}{c'})$$

or

$$f(x) = \frac{g(x)}{G(\frac{1}{c'})}\,e^{-\frac{1}{c'}u(x)} \; .$$

This latter equation, with (4), gives

$$\int_r u(x)f(x)dx = \bar{u} = \frac{1}{G(\frac{1}{c'})} \int_r u(x)g(x)e^{-\frac{1}{c'}u(x)} dx = -\frac{G'(\frac{1}{c'})}{G(\frac{1}{c'})} .$$

By Lemma 1, if the latter has a solution, it has a unique solution. Setting $\frac{1}{c'} = b$, the theorem is proved. #

Lemma 3: Let r and r' be subsets of the real line. Let f(x) and h(y) be continuous functions on r and r', respectively, where h is a one-one mapping of r' to r, with a continuous derivative $h'(y) \neq 0$ on r' and a continuous inverse h^{-1}. Then for a continuous function u(x) defined on r, if

$$f^*(x) = \arg\max_{f(x)} - c_0 - c'\int_r f(x) \ln f(x)\, dx - \int_r f(x)u(x)dx ,$$

then the k*(y) such that

$$k^*(y) = \arg\max_{k(y)} - c_0 - c'\int_{r'} k(y) \ln[\frac{k(y)}{h'(y)}]\, dy - \int_{r'} k(y)u(h(y))dy ,$$

is $k^*(y) = f^*(h(y))h'(y)$.

Proof: Suppose the theorem is false. Then there exists a k**(y) such that

$$-\int_{r'} k^{**}(y)\ln[\frac{k^{**}(y)}{h'(y)}]dy > -\int_{r'} k^*(y)\ln[\frac{k^*(y)}{h'(y)}]dy$$

$$= -\int_{r'} f^*(h(y))h'(y)\ln[f^*(h(y))]dy = -\int_r f^*(x)\ln f^*(x)dx .$$

Now setting $\frac{k^{**}(y)}{h'(y)} = \frac{k^{**}(h^{-1}(x))}{h'(h^{-1}(x))} = f^{**}(x)$, we obtain a contradiction. #

Comment on Lemma 3: Given a minimum information solution for the information definition I(X) and the loss function u(x), we can immediately find a solution for the information definition $I_{h'}(Y)$ and the loss function $u(h(y))$, provided h(y) satisfies the conditions of the lemma.

III. Some Probability Distributions Arising from the Minimum-Information Hypothesis.

The theorems in this section give the minimum-information distributions arising from various forms of the loss function u(x). Theorems 1 and 2 use the information definition I(X) of (2A). The class of continuous distributions arising from the loss function $u(x) = x^{\alpha}, (x \geq 0, \alpha > 0)$, where prices x are measured as deviations from the optimal value $\bar{x}$, are called here "fractional power distributions," and the general equation for their density function is given by Theorem 1. The exponential and one-sided normal distributions are members of this class. These distributions would only arise if a trader always quoted prices that were too high. A related class of continuous distributions are the "fractional-absolute power distributions" of Theorem 2. These result from the loss function $u(x) = |x|^{\alpha}$, (x real, $\alpha > 0$). The double exponential and the normal are members of this class. These distributions arise if the trader's price quotations are symmetrically distributed around $\bar{x}$.

Using the alternative information definition I*(X) of (2B), Theorem 3 describes the distributions arising from the loss function $u(x) = |\ln(x)|^{\alpha}$, $(x \geq 0, \alpha > 0)$, where prices x are now measured as proportions of $\bar{x}$. The lognormal is a member of this class of "log-fractional absolute power distributions." These distributions arise when price quotations are symmetrically high or low as proportions of the optimal value $\bar{x}$.

Theorem 1: In the notation of Lemma 2, let $r = [0,\infty)$ and $u(x) = x^{\alpha}$, where $\alpha > 0$. Then the probability density $f(x)$ which minimizes $c(I) + \overline{u}$, where $I(X) = \int_0^{\infty} f(x) \ln f(x)\, dx$ and $0 < \overline{u} = \int_0^{\infty} u(x)\, f(x)\, dx$ is

$$f(x) = \frac{\alpha}{(c')^{1/\alpha}\Gamma(1/\alpha)} \exp^{-\frac{x^{\alpha}}{c'}},$$

where $c' = \alpha\overline{u}$. (That is, the marginal cost of information is equal to the expected value of the loss function multiplied by α.) Distributions with this density will be referred to as "fractional power distributions." These distributions have as their j-th moment ($j > 0$)

$$m_j = (c')^{j/\alpha} \frac{\Gamma(\frac{j+1}{\alpha})}{\Gamma(1/\alpha)} .$$

Corollary 1.1: For $\alpha = 1$, $f(x)$ is the exponential distribution

$$f(x) = \frac{1}{c'} \exp^{-\frac{x}{c'}}$$

with mean c' and variance c'^2.

Corollary 1.2: For $\alpha = 2$, $f(x)$ is the one-sided normal distribution

$$f(x) = \frac{2}{\sqrt{c'\pi}} \exp^{-\frac{x^2}{c'}}$$

with second moment $\frac{c'}{2}$ mean $\sqrt{\frac{c'}{\pi}}$, and variance $\frac{c'}{2}(1 - 2/\pi)$. (This is the distribution of Brownian motion reflected at the origin.)

Proof of Theorem 1: It will be helpful if we first calculate the integral $\int_0^{\infty} x^k e^{-bx^{\alpha}} dx$, for $k > 0$. Using the substitutions $t = bx^{\alpha}$,

$dt = \alpha b x^{\alpha-1}dx$, and $x^{k-\alpha+1} = \dfrac{t^{\frac{k+1}{\alpha}-1}}{b^{\frac{k+1}{\alpha}-1}}$ we obtain

$$\int_0^\infty x^k e^{-bx^\alpha}dx = \frac{1}{\alpha b^{\frac{k+1}{\alpha}}}\int_0^\infty t^{\frac{k+1}{\alpha}-1}e^{-t}dt = \frac{1}{\alpha b^{\frac{k+1}{\alpha}}}\Gamma\left(\frac{k+1}{\alpha}\right) .$$

Now, in the notation of Lemma 2, we have $G(b) = \int_0^\infty e^{-bx^\alpha}dx = \dfrac{1}{\alpha b^{1/\alpha}}\Gamma(\frac{1}{\alpha})$.

Hence $f(x) = \dfrac{e^{-bx^\alpha}}{\Gamma(\frac{1}{\alpha})/\alpha b^{1/\alpha}}$. Thus $\overline{u} = \int_0^\infty u(x)f(x)dx = \dfrac{\int_0^\infty x^\alpha e^{-bx^\alpha}dx}{\Gamma(\frac{1}{\alpha})/\alpha b^{1/\alpha}} = \dfrac{\Gamma(\frac{\alpha+1}{\alpha})}{b\Gamma(\frac{1}{\alpha})} = \dfrac{1}{b\alpha}$

or $b = \dfrac{1}{\overline{u}\alpha}$. Substituting <u>b</u> back into the expression for $f(x)$, and remembering from Lemma 2 that $b = \frac{1}{c'}$, so that $c' = \alpha\overline{u}$, we obtain Theorem 1. #

<u>Theorem 2</u>: Let r be the real line and $u(x) = |x|^\alpha$, where $\alpha > 0$. Then the probability density $g(x)$ which minimizes $c(I) + \overline{u}$ where

$I(X) = \int_{-\infty}^\infty g(x)\ln g(x)\, dx$ and $0 < \overline{u} = \int_{-\infty}^\infty u(x)g(x)dx$ is

$$g(x) = \tfrac{1}{2} f(|x|)$$

where $f(x)$ is given by Theorem 1. Distributions with the density $g(x)$ will be referred to as "fractional-absolute power distributions." These distributions have as their j-th <u>absolute</u> moment ($j \geq 0$)

$$m_j = (c')^{j/\alpha}\frac{\Gamma(\frac{j+1}{\alpha})}{\Gamma(1/\alpha)} .$$

<u>Corollary 2.1</u>: For $\alpha = 1$, $g(x)$ is the double exponential distribution

$$g(x) = \frac{1}{2c'}\exp^{-\frac{|x|}{c'}}$$

with first absolute moment $\overline{u} = c'$, mean 0, and variance $2\overline{u}^2 = 2c'^2$

Corollary 2.2: For $\alpha = 2$, g(x) is the normal distribution with variance $\frac{c'}{2} = \overline{u}$ and mean 0.

Proof of Theorem 2: From the proof of Theorem 1 and by symmetry we have $\int_{-\infty}^{\infty} |x|^k e^{-b|x|^\alpha} dx = \frac{2}{\alpha b^{\frac{k+1}{\alpha}}} \Gamma(\frac{k+1}{\alpha})$. Thus $G(b) = \frac{2}{\alpha b^{1/\alpha}} \Gamma(\frac{1}{\alpha})$ and

$f(x) = \frac{e^{-b|x|^\alpha}}{2\Gamma(\frac{1}{\alpha})/\alpha b^{1/\alpha}}$. Solving for $\overline{u} = \int_{-\infty}^{\infty} |x|^\alpha f(x)dx$ we find that $\overline{u}$ and thus b have the same values as in the proof of Theorem 1. Substituting the value for b into f(x), we obtain Theorem 2. #

Theorem 3: Here we consider x measured as a proportion of its optimal value $\overline{x}$. Let $r = [0, \infty)$ and $u(x) = |\ln(x)|^\alpha$, where $\alpha > 0$. Then the probability density function which minimizes $c(I^*) + \overline{u}$, where $I^*(X) = \int_0^\infty f(x)\ln[xf(x)]dx$ and $\overline{u} = \int_0^\infty u(x)f(x)dx$ is

$$f(x) = \frac{\alpha}{2} \frac{1}{(c')^{1/\alpha}} \frac{1}{\Gamma(\frac{1}{\alpha})} \frac{1}{x} \exp^{-\frac{|\ln x|^\alpha}{c'}}$$

where $c' = \alpha\overline{u}$. Distributions with this density will be called "log-fractional absolute power distributions."

Corollary 3.1: When $\alpha = 2$, f(x) is the lognormal distribution

$$f(x) = \frac{1}{\sqrt{c'\pi}} \frac{1}{x} \exp^{-\frac{(\ln x)^2}{c'}}$$

with mean 0, and where $c'/2 = \overline{u}$ is the variance of ln x.

Proof of Theorem 3: The function $h(y) = \ln y$ on $y \in [0, \infty)$ is continuous, with a derivative $h'(y) = \frac{1}{y} \neq 0$, and maps $[0, \infty)$ to the entire real line. It has a continuous inverse $h^{-1}(x) = e^x$, x real. Hence Theorem 3 follows from Theorem 2, by application of Lemma 3. #

IV. Leptokurtosis in the Distribution of First-Differences in Log Prices.

One fairly well established empirical fact is that in many speculative markets log differences in prices are roughly symmetric, but exhibit a marked degree of leptokurtosis.[2] In one example, the kurtosis parameter (K) for the distribution of log changes in daily exchange rate data was found to have a value of about K = 10.[3] This contrasts with the normal distribution where K has the value K = 3.

One possible explanation is found in Theorem 2. If we consider this class of minimum-information distributions, of which the normal is one member, the kurtosis parameter is easily calculated to be

$$K = \frac{\Gamma(5/\alpha)\Gamma(1/\alpha)}{[\Gamma(3/\alpha)]^2} .$$

Values of K for selected values of α are given below:

α	K
4	2.2
3	2.4
2	3
1	6
2/3	12.3
1/2	25.2

The normal is the case α = 2. For values of α greater than 2, the minimum-information distribution is platykurtic, while for values less than 2, the minimum-information distribution is leptokurtic. The previously mentioned empirical finding of K = 10 corresponds to a value of α less than 1.

If log changes in prices correspond to a Theorem 2 distribution, then relative price levels correspond to a Theorem 3 distribution. The associated loss function $u(x) = |\ln x|^{\alpha}$ is symmetric about x=1. Thus, viewing x as the price announced by a trader, where x is measured as a proportion of the optimal price $\overline{x}$, this class of distributions corresponds to the case the loss function assigns equal weights to proportional departures from $\overline{x}$, whether up or down. The rate at which weight accumulates in either direction from x=1 is governed by α. As seen in the table below, for α=1, the loss function u(x) increases from 1 to 8, and from 1 to 1/8, in uniform increments of .693. For α>1, u(x) gives proportional increments in x more weight the further away

	x	1/8	1/4	1/2	1	2	4	8
u(x)	$\lvert\ln x\rvert^2$	4.324	1.922	.480	0	.480	1.922	4.324
	$\lvert\ln x\rvert$	2.079	1.386	.693	0	.693	1.386	2.079
	$\lvert\ln x\rvert^{1/2}$	1.442	1.177	.833	0	.833	1.177	1.442

from the optimal price of x=1, while for α<1, proportional increments are given less weight the further away from the optimal price of x=1. Intuitively, we might expect the loss functional to give less weight to proportional increments that are further out, and hence α to be less than 1. Why? Simply because if the price is much too low or too high, the trader will be hit almost surely <u>only</u> on his bid side or his ask side, and the trading volume will be such that his losses will not be proportionately magnified by an even higher or lower price. The distribution is then necessarily leptokurtic. In any case, if we restrict our attention to the family of

Theorem 3 of which the lognormal is a member, leptokurtosis in the distribution of log changes will result if $\alpha < 2$; that is, if the loss functional is less convex than the quadratic.

V. Comment and Conclusion.

This paper has described a rational theory of price distributions in a speculative market such as that for foreign exchange. The distribution of such prices is not assigned by God or nature from outside the system. Rather, these prices are announced by traders who use rational criteria in making their decisions. The procedure for doing so was modelled here as the selection of a probability distribution from which price drawings are made.

Conceptually there is an optimal price $\overline{x}$, unknown to the trader, such that, if it were announced at the current time, would maximize some objective function (making money, say). The hypothesized existence of this objective function implies the trader faces a loss functional, a function of the announced price, which represents departure from the optimum whenever the announced price is not the optimal price. In addition, it is assumed that the trader can gain information about the location of this optimal price, and the more information the trader has will be reflected in the shape of the probability distribution from which prices are drawn.

For the loss functionals studied in this paper, the trader will acquire information up to the point where the marginal cost of the information c' is a scalar multiple of the expected loss $\overline{u}$: $c' = \alpha\overline{u}$, where α is a parameter governing the steepness of the tails of the convex loss functional. For example, it was seen that for the loss functional $u(x) = |\ln(x)|^{\alpha}$, where x is the announced price expressed as a proportion of the optimal price, and for $\alpha = 2$, the resulting distribution was the lognormal distribution. If $\alpha > 2$, the resulting distribution is platykurtic, while for $\alpha < 2$, it is leptokurtic. In addition, the variance of $\ln(x)$ is determined precisely by the marginal cost of information, c'. Thus, if the marginal cost of information changes over time, this implies immediately a time-varying variance. And if the marginal cost of information is low, the distribution will be concentrated around the optimal price $\overline{x}$.

In conclusion, then, we see that the dispersion of price announcements, as measured by the variance, is governed by the marginal cost of information. But the kurtosis--the allocation of weight to the center and tails of the distribution, as compared to the intermediate ranges--is governed by the convexity of the loss functional. Thus the fundamental shape of price distributions in speculative markets would seem to be determined by simple, rational choice.

Footnotes

1. By convention, $p_i \ln p_i = 0$ when $p_i = 0$.

2. Extensive references are cited in Grabbe (1981).

3. In Grabbe (1981). Daily changes in the logarithm of the dollar-DM spot exchange rate over the period July 1973 to June 1979 had a sample kurtosis of 11.88. For one-month forward rates, the value was 9.64.

Bibliography

1. Barbosa, F. de H. (1975), Rational Random Behavior: Extentions and Applications, Ph.D. dissertation, The University of Chicago.

2. Campbell, L. L. (1972), "Characterization of Entropy of Probability Distributions on the Real Line," Information and Control, 21, 329-338.

3. Dowson, D. C. and A. Wragg (1973), "Minimum-Entropy Distributions Having Prescribed First and Second Moments," IEEE Trans. Inform. Theory, September, 689-693.

4. Everett, III, Hugh (1973), "The Theory of the Universal Wave Function," in The Many-Worlds Interpretation of Quantum Mechanics, edited by Bryce S. DeWitt and Neill Graham, Princeton University Press, Princeton.

5. Grabbe, J. Orlin (1981), "Are Forward Exchange Rates Really Useful Predictors of Future Spot Rates?--Some Evidence from Daily Dollar-DM Data," mimeo, Wharton School, University of Pennsylvania, November.

6. Guiasu, Silviu (1977). Information Theory with Applications, McGraw Hill, New York.

7. Ingarden, R. S. and Kossakowski, A. (1971), "Poisson Probability Distribution and Information Thermodynamics," Bull. Acad. Sci. Polonaise, 19, 83-86.

8. Kolmogorov, Andrei N. (1956), "On the Shannon Theory of Information Transmission in the Case of Continuous Signals," IEEE Trans. Inform. Theory, vol. IT-2, 102-108.

9. Kullback, S. and Leibler, R. A. (1951). "On Information and Sufficiency," Annals of Mathematical Statistics, 22, 79-86.

10. Shannon, Claude E. (1948). "A Mathematical Theory of Communication," Bell System Technical Journal, 27, 379-423 and 623-656.

11. Theil, Henri (1980), The System-Wide Approach to Microeconomics, The University of Chicago Press, Chicago.

12. Tribus, M. (1969), Rational Descriptions, Decisions, and Design, Pergamon Press, New York.

13. Wiener, Norbert (1961), Cybernetics, 2nd edition, The MIT Press, Cambridge.