PART IV: THE 1963 MODEL OF PRICE CHANGE

My efforts to improve on Bachelier's Brownian *model started with markets on which the dominant factor is the highly nonGaussian nature of the distribution's tails. In IBM Report NC-87,* "The Variation of certain speculative prices, *published on March 26, 1962, the title pointedly meant* not necessarily all. *For external publication, the hefty NC-87 was split. Its core became M 1963b{E14}, which is the centerpiece of this part. My interests having changed, what was left of NC-87 appeared years later, as M 1967b{E15}; this awkward composite adds some data on cotton and continues with wheat, railroad securities and interest rates, but also includes answers to criticism of M 1963b{E14} and other material that made the text closer to being self-contained. A final small piece from NC-87 is published for the first time as Pre-publication Appendix I to Chapter E14.*

&&&&&&&&&&&&&&&&&&&&&&&&&&

The Journal of Business, 36, 1963, 394-419 & 45, 1972, 542-3
Econometrica, 31, 1963, 757-758. Current Contents, 14, 1982, 20.

E14

The variation of certain speculative prices

✦ **Pre-publication abstract (M 1962a).** The classic model of the temporal variation of speculative prices (Bachelier 1900) assumes that successive changes of a price $Z(t)$ are independent Gaussian random variables. But, even if $Z(t)$ is replaced by $\log Z(t)$, this model is contradicted by facts in four ways, at least:

(1) Large price changes are *much more frequent* than predicted by the Gaussian; this reflects the "excessively peaked" ("leptokurtic") character of price relatives, which has been well-established since at least 1915.

(2) Large practically instantaneous price changes *occur often*, contrary to prediction, and it seems that they must be explained by causal rather than stochastic models.

(3) Successive price changes *do not* "look" independent, but rather exhibit a large number of recognizable patterns, which are, of course, the basis of the technical analysis of stocks.

(4) Price records *do not* look stationary, and statistical expressions such as the sample variance take very different values at different times; this nonstationarity seems to put a precise statistical model of price change out of the question.

I shall show that there is a simple way to solve difficulties (1), (2) and (4), and – to some extent – difficulty (3). This will imply that it is *not* necessary to give up the stationary stochastic models. Suppose indeed that the price relatives are so extremely leptokurtic (1), as to lead to infinite values for the population variance, and for other population moments beyond the first. This could – and indeed does – explain the erratic behavior of the sample moments (4), and the sample paths generated by such models would indeed by expected to include large discontinuities (2). Additionally, some features of the dependence between successive changes (3) could be taken into account by injecting a comparatively limited weakening asymptotic? of the hypothesis of independence; that is, "patterns" that have such a small probability in a Gaussian function that their occurrence by chance is practically impossible, now acquire a credibly large probability of occurring by chance.

As known in the case of the Cauchy distribution, having an infinite variance does not prevent a distribution from being quite proper, but it does make it quite peculiar. For example, the classical central limit theorem is inapplicable, and the largest of M addends is not negligibly small but rather provides an appreciable proportion of their sum. Fortunately, these peculiar consequences actually happen to describe certain well-known features of the behavior of prices.

The basic distribution with an infinite variance is scaling with an exponent between 1 and 2. My theory of prices is based upon distributions with two scaling tails, as well as upon L-stable distributions. The latter are akin to the scaling law, and appear in the first significant generalization of the classical central limit theorem. My theory is related to my earlier work on the distribution of personal income. ✦

I. INTRODUCTION

Louis Bachelier is a name mentioned in relation to diffusion processes in physics. Until very recently, however, few people realized that his path-breaking contribution, Bachelier 1900, was a by-product of the construction of a random-walk model for security and commodity markets. Let $Z(t)$ be the price of a stock, or of a unit of a commodity, at the end of time period t. Then, Bachelier's simplest and most important model assumes that successive differences of the form $Z(t+T)-Z(t)$ are independent Gaussian random variables with zero mean and with variance proportional to the differencing interval T.

That simplest model implicitly assumes that the variance of the differences $Z(t+T)-Z(t)$ is independent of the level of $Z(t)$. There is reason to expect, however, that the standard deviation of $\Delta Z(t)$ will be proportional to the price level, which is why many authors suggest that the original assumption of independent increments of $Z(t)$ be replaced by the assumption of independent and Gaussian increments of $\log_e Z(t)$.

Despite the fundamental importance of Bachelier's process, which has come to be called "Brownian motion," it is now obvious that it does not account for the abundant data accumulated since 1900 by empirical economists. Simply stated, *the empirical distributions of price changes are usually too "peaked" to be viewed as samples from Gaussian populations.* To the best of my knowledge, the first to note this fact was Mitchell 1915. But unquestionable proof was only given by Olivier 1926 and Mills 1927. Other evidence, regarding either $Z(t)$ or $\log Z(t)$, can be found in Larson 1960, Osborne 1959 and Alexander 1961.

That is, the histograms of price changes are indeed unimodal and their central "bells" are reminiscent of the "Gaussian ogive." But there are typically so many "outliers" that ogives fitted to the mean square of price changes are much lower and flatter than the distribution of the data themselves (see, Fig. 1). The tails of the distributions of price changes are in fact so extraordinarily long that the sample second moments typically vary in an erratic fashion. For example, the second moment reproduced in Figure 2 does not seem to tend to any limit even though the sample size is enormous by economic standards.

It is my opinion that these facts warrant a radically new approach to the problem of price variation in speculative markets. The purpose of this paper will be to present and test a new model that incorporates this belief. (A closely related approach has also proved successful in other contexts; see M 1963e{E3}. But I believe that each of the applications should stand on their own feet and I have minimized the number of cross references.

The model I propose begins like the Bachelier process as applied to $\log_e Z(t)$ instead of $Z(t)$. The major change is that I replace the Gaussian distribution throughout by "L-stable," probability laws which were first described in Lévy 1925. In a somewhat complex way, the Gaussian is a limiting case of this new family, so the new model is actually a generalization of that of Bachelier.

Since the L-stable probability laws are relatively unknown, I shall begin with a discussion of some of the more important mathematical properties of these laws. Following this, the results of empirical tests of the L-stable model will be examined. The remaining sections of the paper will

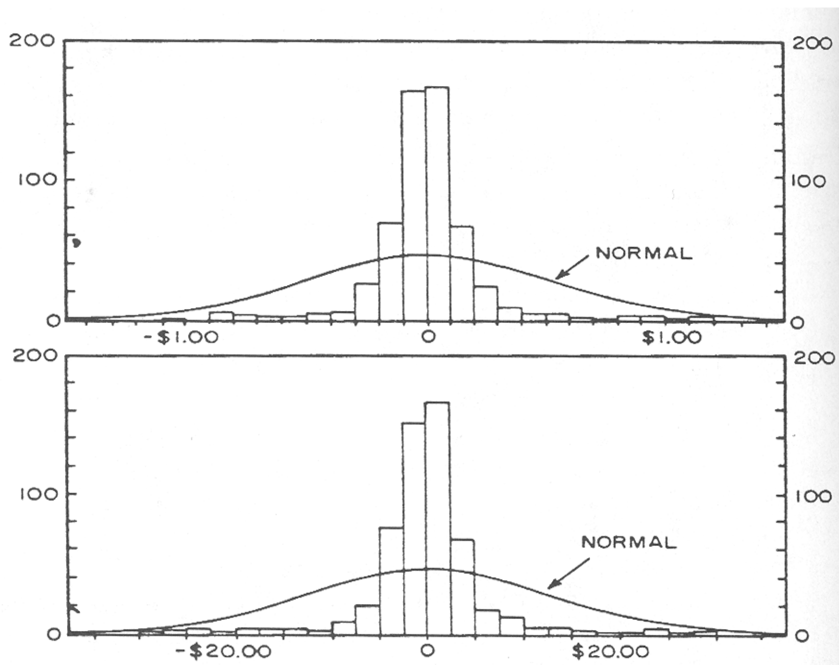

FIGURE C14-1. Two histograms illustrating departure from normality of the fifth and tenth difference of monthly wool prices, 1890-1937. In each case, the continuous bell-shaped curve represents the Gaussian "interpolate" from -3σ to 3σ based upon the sample variance. Source: Tintner 1940.

then be devoted to a discussion of some of the more sophisticated mathematical and descriptive properties of the L-stable model. I shall, in particular, examine its bearing on the very possibility of implementing the stop-loss rules of speculation.

II. MATHEMATICAL TOOLS: L-STABLE DISTRIBUTIONS

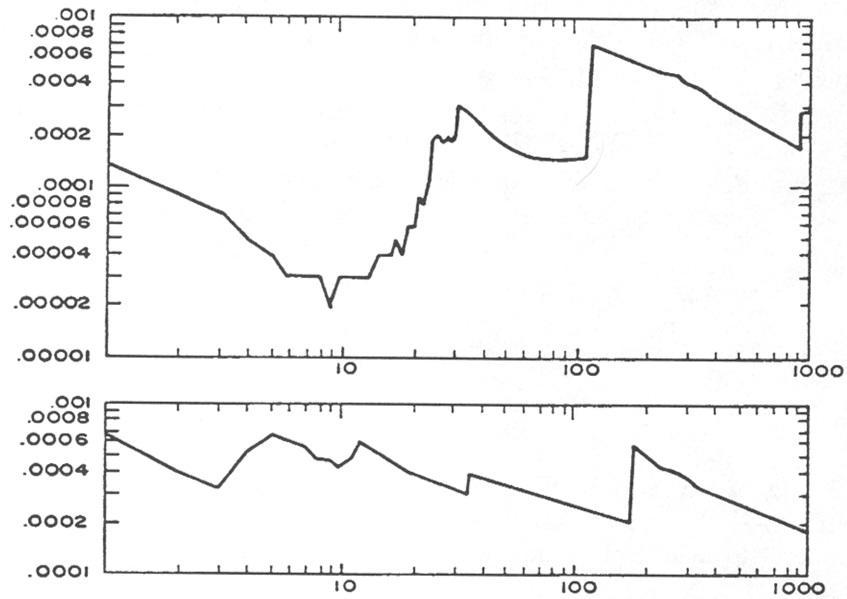

FIGURE C14-2. Both graphs represent the sequential variation of the sample second moment of cotton price changes. The horizontal scale represents time in days, with two different origins T_0. On the upper graph, T_0 was September 21, 1900; on the lower graph, T_0 was August 1, 1900. The vertical lines represent the value of the function

$$(T - T_0)^{-1} \sum_{t = T_0}^{t = T} [L(t, 1)]^2,$$

where $L(t, 1) = \log_e Z(t+1) - \log_e Z(t)$ and $Z(t)$ is the closing spot price of cotton on day t. I am grateful to the United States Department of Agriculture for making these data available.

II.A. "L-stability" of the Gaussian distribution and generalization of the concept of L-stability

One of the principal attractions of the modified Bachelier process is that the logarithmic relative

$$L(t, T) = \log_e Z(t + T) - \log_e Z(t),$$

is a Gaussian random variable for *every* value of T; the only thing that changes with T is the standard deviation of $L(t, T)$. This feature is the consequence of the following fact:

Let G' and G'' be two independent Gaussian random variables, of zero means and of mean squares equal to σ'^2 and σ''^2, respectively. Then the sum $G' + G''$ is also a Gaussian variable of mean square equal to $\sigma'^2 + \sigma''^2$. In particular, the "reduced" Gaussian variable, with zero mean and unit square, is a solution to

$$s'U + s''U = sU, \tag{S}$$

where s is a function of s' and s'' given by the auxiliary relation

$$s^2 = s'^2 + s''^2. \tag{A$_2$}$$

It should be stressed that, from the viewpoint of the equation (S) and relation A_2, the quantities s', s'', and s are simply scale factors that "happen" to be closely related to the root-mean-square in the Gaussian case.

The property (S) expresses a kind of L-stability or invariance under addition, which is so fundamental in probability theory that it came to be referred to simply as *L-stability*. The Gaussian is the only solution of equation (S) for which the second moment is finite – or for which the relation A_2 is satisfied. When the variance is allowed to be infinite, however, (S) possesses many other solutions. This was shown constructively by Cauchy, who considered the random variable U for which

$$\Pr\{U > u\} = \Pr\{U < -u\} = 1/2 - (1/\pi)\tan^{-1} u,$$

so that its density is of the form

$$d \Pr\{U < u\} = \frac{1}{\pi(1 + u^2)}.$$

For this law, integral moments of all orders are infinite, and the auxiliary relation takes the form

$$(\mathrm{A}_1) \qquad s = s' + s'',$$

where the scale factors s', s'', and s are not defined by any moment.

The general solution of equation (S) was discovered by Lévy 1925. (The most accessible source on these problems is, however, Gnedenko & Kolmogorov 1954.) The logarithm of its characteristic function takes the form

$$(\mathrm{L}) \qquad \log \int_{-\infty}^{\infty} \exp(iuz) d \Pr\{U < u\} = i\delta z - \gamma |z|^{\alpha} \left\{ 1 + \frac{i\beta z}{|z|} \tan \frac{\alpha\pi}{2} \right\}.$$

It is clear that the Gaussian law and the law of Cauchy are stable and that they correspond to the cases ($\alpha = 2$; β arbitrary) and ($\alpha = 1$; $\beta = 0$), respectively.

Equation (L) determines a family of distribution and density functions $\Pr\{U < u\}$ and $d \Pr\{U < u\}$ that depend continuously upon four parameters. These four parameters also happen to play the roles the Pearson classification associates with the first four moments of U.

First of all, the α is an index of "peakedness" that varies in $]0, 2]$, that is, from 0 (excluded) to 2 (included). This α will turn out to be intimately related to the scaling exponent. The β is an index of "skewness" that can vary from -1 to $+1$, except that, if $\alpha = 1$, β must vanish. If $\beta = 0$, the stable densities are symmetric.

One can say that α and β together determine the "type" of a stable random variable. Such a variable can be called "reduced" if $\gamma = 1$ and $\delta = 0$. It is easy to see that, if U is reduced, sU is a stable variable with the same α, β and δ, and γ equal to s^{α}. This means that the third parameter, γ, is a scale factor raised to the power of α. Suppose now that U' and U'' are two independent stable variables, reduced and having the same values for α and β. It is well-known that the characteristic function of $s'U' + s''U''$ is the product of those of $s'U'$ and of $s''U''$. Therefore, the equation (S) is readily seen to be accompanied by the auxiliary relation

$$(\mathrm{A}) \qquad s^{\alpha} = s'^{\alpha} + s''^{\alpha}.$$

More generally, suppose that U' and U'' are stable, have the same values of α, β and of $\delta = 0$, but have different values of γ (respectively, γ' and γ''), the sum $U' + U''$ is stable and has the parameters α, β, $\gamma = \gamma' + \gamma''$ and $\delta = 0$. Now recall the familiar property of the Gaussian distribution, that when two Gaussian variables are added, one must add their " variances." The variance is a mean-square and is the square of a scale factor. The role of a scale factor is now played by γ, and that of a variance by a scale factor raised to the power α.

The final parameter is δ; strictly speaking, equation (S) requires that $\delta = 0$, but we have added the term $i\delta z$ to (PL) in order to introduce a location parameter. If $1 < \alpha \leq 2$, so that $E(U)$ is finite, one has $\delta = E(U)$. If $\beta = 0$, so that the stable variable has a symmetric density function, δ is the median or modal value of U. But when $0 < \alpha < 1$, with $\beta \neq 0$, δ has no obvious interpretation.

II.B. Addition of more than two stable random variables

Let the independent variables U_n satisfy the condition (PL) with values of α, β, γ, and δ equal for all n. The logarithm of the characteristic function of

$$S_N = U_1 + U_2 + \ldots U_n + \ldots U_N$$

is N times the logarithm of the characteristic function of U_n, and equals

$$i\,\delta Nz - N\gamma |z|^{\alpha}[1 + i\beta(z/|z|)\tan(\alpha\pi/2)].$$

Thus S_N is stable with the same α and β as U_n, and with parameters δ and γ multiplied by N. It readily follows that

$$U_n - \delta \text{ and } N^{-1/\alpha}\sum_{n=1}^{N} U_n - \delta$$

have identical characteristic functions and thus are identically distributed random variables. (This is, or course, a most familiar fact in the Gaussian case, $\alpha = 2$.)

The generalization of the classical "$T^{1/2}$ law." In the Gaussian model of Bachelier, in which daily increments of $Z(t)$ are Gaussian with the standard deviation $\sigma(1)$, the standard deviation of $\Delta Z(t)$, where Δ is taken over T days, is equal to $\sigma(T) = T^{1/2}\sigma(1)$.

The corresponding prediction of my model is as follows: Consider any scale factor such as the intersextile range, that is, the difference between the quantity U^+ which is exceeded by one-sixth of the data, and the quantity U^- which is larger than one-sixth of the data. It is easily found that the expected range satisfies

$$E[U^+(T) - U^-(T)] = T^{1/\alpha}E[U^+(1) - U^-(1)].$$

We should also expect that the deviations from these expectations exceed those observed in the Gaussian case.

Differences between successive means of $Z(t)$. In all cases, the average of $Z(t)$, taken over the time span $t^0 + 1$ to $t^0 + N$, can be written as:

$$(1/N)[Z(t^0+1) + Z(t^0+2) + ...Z(t^0+N)]$$
$$= (1/N)\{N\,Z(t^0+1) + (N-1)[Z(t^0+2) - Z(t^0+1)] + ...$$
$$+ (N-n)[Z(t^0+n+1) - Z(t^0+n)] + ...[Z(t^0+N) - Z(t^0+N-1)]\}.$$

To the contrary, let the average over the time span $t^0 - N + 1$ to t^0 be written as

$$(1/N)\{N\,Z(t^0) + (N-1)[Z(t^0) - Z(t^0-1)]...$$
$$+ (N-n)[Z(t^0-n+1) - Z(t^0-n)]...$$
$$+ [Z(t^0-N+2) - Z(t^0-N+1)]\}.$$

Thus, if the expression $Z(t+1) - Z(t)$ is a stable variable $U(t)$ with $\delta = 0$, the difference between successive means of values of Z is given by

$$U(t^0) + [(N-1)/N][U(t^0+1) + U(t^0-1)]$$
$$+ ...[(N-n)/N][U(t^0+n) + U(t^0-n)]$$
$$+ ...[U(t^0+N-1)...U(t^0-N+1)].$$

This is clearly a stable variable, with the same α and β as the original U, and with a scale parameter equal to

$$\gamma^0(N) = [1 + 2(N-1)^\alpha N^{-\alpha} + ...2(N-n)^\alpha N^{-\alpha} + ... + 2]\gamma(U).$$

As $N \to \infty$, one has

$$\frac{\gamma^0(N)}{\gamma(U)} \to \frac{2N}{(\alpha+1)},$$

whereas a genuine monthly change of $Z(t)$ has a parameter $\gamma(N) = N\gamma(U)$. Thus, the effect of averaging is to multiply γ by the expression $2/(\alpha+1)$, which is smaller than 1 if $\alpha > 1$.

III.C. L-stable distributions and scaling

Except for the Gaussian limit case, the densities of the stable random variables follow a generalization of the asymptotic behavior of the Cauchy law. It is clear, for example, that as $u \to \infty$, the Cauchy density behaves as follows:

$$u \Pr\{U > u\} = u \Pr\{U < -u\} \to 1/\pi.$$

More generally, Lévy has shown that the tails of *all* nonGaussian stable laws follow an asymptotic form of scaling. There exist two constants, $C' = \sigma'^\alpha$ and $C'' = \sigma''^\alpha$, linked by $\beta = (C' - C'')/(C' + C'')$, such that,

$$\text{when } u \to \infty,\ u^\alpha \Pr\{U > u\} \to C' = \sigma'^\alpha \text{ and } u^\alpha \Pr\{U < -u\} \to C'' = \sigma''^\alpha.$$

Hence, *both* tails are scaling if $|\beta| \neq 1$, a solid reason for replacing the term "stable nonGaussian" by the less negative one of *"L-stable."* The two numbers σ' and σ'' share the role of the standard deviation of a Gaussian variable. They will be denoted as the "standard positive deviation" and the "standard negative deviation," respectively.

Now consider the two extreme cases: when $\beta = 1$, hence $C'' = 0$, and when $\beta = -1$, hence $C' = 0$). In those cases, one of the tails (negative and positive, respectively) decreases faster than the scaling distribution of index α. In fact, one can prove (Skorohod 1954-1961) that the short tail withers away even faster that the Gaussian density so that the extreme cases of stable laws are, for all practical purposes, J-shaped. They play an important role in my theory of the distributions of personal income and of city sizes. A number of further properties of L-stable laws may therefore be found in my publications devoted to these topics. See M 1960i{E10}, 1963p{E11} and 1962g{E12}.

II.D. The L-stable variables as the only possible limits of weighted sums of independent, identically distributed addends

The L-stability of the Gaussian law can be considered to be only a matter of convenience, and it often thought that the following property is more important.

Let the U_n be independent, identically distributed random variables, with a finite $\sigma^2 = E[U_n - E(U)]^2$. Then the classical central limit theorem asserts that

$$\lim_{N \to \infty} N^{-1/2} \sigma^{-1} \sum_{n=1}^{N} [U_n - E(U)]$$

is a reduced Gaussian variable.

This result is, of course, the basis of the explanation of the presumed occurrence of the Gaussian law in many practical applications relative to sums of a variety of random effects. But the essential thing in all these aggregative arguments is not that $\Sigma[U_n - E(U)]$ is weighted by any special factor, such as $N^{-1/2}$, but rather that the following is true:

There exist two functions, $A(N)$ and $B(N)$, such that, as $N \to \infty$, the weighted sum

(L) $$A(N) \sum_{n=1}^{N} U_n - B(N),$$

has a limit that is finite and is not reduced to a nonrandom constant.

If the variance of U_n is not finite, however, condition (L) may remain satisfied while the limit ceases to be Gaussian. For example, if U_n is stable nonGaussian, the linearly weighted sum

$$N^{-1/\alpha} \sum (U_n - \delta)$$

was seen to be *identical in law* to U_n, so that the "limit" of that expression is already attained for $N = 1$ and a stable nonGaussian law. Let us now suppose that U_n is asymptotically scaling with $0 < \alpha < 2$, but not stable. Then the limit exists, and it follows the L-stable law having the same value of α. As in the L-stability argument, the function $A(N)$ can be chosen equal to $N^{-1/\alpha}$. These results are crucial but I had better not attempt to rederive them here. The full mathematical argument is available in the literature. I have constructed various heuristic arguments to buttress it. But experience shows that an argument intended to be illuminating often comes across as basing far-reaching conclusions on loose thoughts. Let me therefore just quote the facts:

The Doeblin-Gnedenko conditions. The problem of the existence of a limit for $A(N)\Sigma U_n - B(N)$ can be solved by introducing the following generalization of asymptotic scaling (Gnedenko & Kolmogorov 1954). Introduce the notations

$$\Pr\{U > u\} = Q'(u)u^{-\alpha}; \; \Pr\{U < -u\} = Q''(u)u^{-\alpha}.$$

The term *Doeblin-Gnedenko condition* will denote the following statements: (a) when $u \to \infty$, $Q'(u)/Q''(u)$ tends to a limit C'/C''; (b) there exists a value of $\alpha > 0$ such that for every $k > 0$, and for $u \to \infty$, one has

$$\frac{Q'(u) + Q''(u)}{Q'(ku) + Q''(ku)} \to 1.$$

These conditions generalize the scaling distribution, for which $Q'(u)$ and $Q''(u)$ themselves tend to limits as $u \to \infty$. With their help, and unless $\alpha = 1$, the problem of the existence of weighting factors $A(N)$ and $B(N)$ is solved by the following theorem:

If the U_n are independent, identically distributed random variables, there may exist no functions $A(N)$ and $B(N)$ such that $A(N)\Sigma U_n - B(N)$ tends to a proper limit. But, if such functions $A(N)$ and $B(N)$ exist, one knows that the limit is one of the solutions of the L-stability equation (S). *More precisely, the limit is Gaussian if, and only if, the U_n has finite variance; the limit is nonGaussian if, and only if, the Doeblin-Gnedenko conditions are satisfied for some* $0 < \alpha < 2$. *Then,* $\beta = (C' - C'')/(C' + C'')$ *and $A(N)$ is determined by the requirement that*

$$N \Pr\{U > uA^{-1}(N)\} \to C'u^{-\alpha}.$$

(For all values of α, the Doeblin-Gnedenko condition (b) also plays a central role in the study of the distribution of the random variable $\max U_n$.)

As an application of the above definition and theorem, let us examine the product of two independent, identically distributed scaling (but not stable) variables U' and U''. First of all, for $u > 0$, one can write

$$\Pr\{U'U'' > u\} = \Pr\{U' > 0;\ U'' > 0;\ \text{and } \log U' + \log U'' > \log u\}$$
$$+ \Pr\{U' < 0;\ U'' < 0;\ \text{and } \log |U'| + \log |U''| > \log u\}.$$

But it follows from the scaling distribution that

$$\Pr\{U > e^z\} \sim C' \exp(-\alpha z) \text{ and } \Pr\{U < -e^z\} \sim C'' \exp(-\alpha z),$$

where U is either U' or U''. Hence, the two terms P' and P'' that add up to $\Pr\{U'U'' > u\}$ satisfy

$$P'C'^2 \alpha z \exp(-\alpha z) \text{ and } P''C''^2 \alpha z \exp(-\alpha z).$$

Therefore,

$$\Pr\{U'U'' > u\} \sim \alpha(C'^2 + C''^2)(\log_e u)u^{-\alpha}.$$

Similarly,

$$\Pr\{U'U'' < -u\} \sim \alpha 2C'C''(\log_e u)u^{-\alpha}.$$

It is obvious that the Doeblin-Gnedenko conditions are satisfied for the functions $Q'(u) \sim (C'^2 + C''^2)\alpha \log_e u$ and $Q''(u) \sim 2C'C''\alpha \log_e u$. Hence the weighted expression

$$(N \log N)^{-1/\alpha} \sum_{n=1}^{N} U'_n U''_n$$

converges toward a L-stable limit with the exponent α and the skewness

$$\beta = \frac{C'^2 + C''^2 - 2C'C''}{C'^2 + C''^2 + 2C'C''} = \left[\frac{C' - C''}{C' + C''} \right]^2 \geq 0.$$

In particular, the positive tail should always be bigger than the negative tail.

II.E. Shape of the L-stable distributions outside the asymptotic range

There are closed expressions for three cases of L-stable densities: Gauss ($\alpha = 2, \beta = 0$), Cauchy ($\alpha = 1, \beta - 0$), and third, ($\alpha = 1/2; \beta = 1$). In every other case, we only know the following: (a) the densities are always unimodal; (b) the densities depend continuously upon the parameters; (c) if $\beta > 0$, the positive tail is the fatter – hence, if the mean is finite (i.e., if $1 < \alpha < 2$), it is greater than the median.

To go further, I had to resort to numerical calculations. Let us, however, begin by interpolative arguments.

The symmetric cases, $\beta = 0$. For $\alpha = 1$, one has the Cauchy density $[\pi(1 + u^2)]^{-1}$. It is always *smaller* than the scaling density $1/\pi u^2$ toward which it converges as $u \to \infty$. Therefore, $\Pr\{U > u\} < 1/\pi\, u$, and it follows that for $\alpha = 1$, the doubly logarithmic graph of $\log_e[\Pr\{U > u\}]$ is entirely on the left side of its straight asymptote. By continuity, the same shape must appear when α is only a little higher or a little lower than 1.

For $\alpha = 2$, the doubly logarithmic graph of the Gaussian $\log_e \Pr\{(U > u)\}$ drops very quickly to negligible values. Hence, again by continuity, the graph must also begin with decreasing rapidly when α is just below 2. But, since its ultimate slope is close to 2, it must have a point of inflection corresponding to a maximum slope greater than 2, and it must begin by "overshooting" its straight asymptote.

Interpolating between 1 and 2, we see that there exists a smallest value of α, call it $\tilde{\alpha}$, for which the doubly logarithmic graph begins by overshooting its asymptote. In the neighborhood of $\tilde{\alpha}$, the asymptotic α can be measured as a slope even if the sample is small. If $\alpha < \tilde{\alpha}$, the asymptotic slope will be underestimated by the slope of small samples; for $\alpha > \tilde{\alpha}$ it will be overestimated. The numerical evaluation of the densities yields a value of $\tilde{\alpha}$ in the neighborhood of 1.5. A graphical presentation of the results of this section is given in Figure 3.

The skew cases. If the positive tail is fatter than the negative one, it may well happen that the doubly logarithmic graph of the positive tail begins by overshooting its asymptote, while the doubly logarithmic graph of the negative tail does not. Hence, there are two critical values of α^0, one for each tail. If the skewness is slight, if α lies between the critical values, and if the sample size is not large enough, then the graphs of the two tails will have slightly different overall apparent slopes.

II.F. Joint distribution of independent L-stable variables

Let $p_1(u_1)$ and $p_2(u_2)$ be the densities of U_1 and of U_2. If both u_1 and u_2 are large, the joint probability density is given by

$$p^0(u_1, u_2) = \alpha C'_1 u_1^{-(\alpha+1)} \alpha C'_2 u_2^{-(\alpha+1)} = \alpha^2 C'_1 C'_2 (u_1 u_2)^{-(\alpha+1)}.$$

The lines of equal probability belong to hyperbolas $u_1 u_2 = \text{constant}$. They link together as in Figure 4, into fattened signs +. Near their maxima, $\log_e p_1(u_1)$ and $\log_e p_2(u_2)$ are approximated by $\alpha_1 - (u_1/b_1)^2$ and $\alpha_2 - (u_2/b_2)^2$. Hence, the probability isolines are of the form

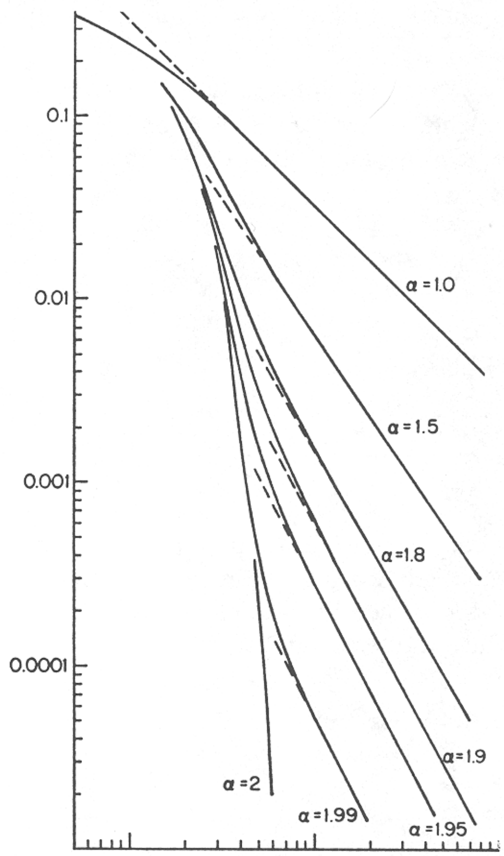

FIGURE C14-3. The various lines are doubly logarithmic plots of the symmetric L-stable probability distributions with $\delta = 0$, $\gamma = 1$, $\beta = 0$ and α as marked. Horizontally: $\log_e u$; vertically: $\log_e \Pr\{U > u\} = \log_e \Pr\{U < -u\}$. Sources: unpublished tables based upon numerical computations performed at the author's request by the IBM T. J. Watson Research Center.

$$(u_1/b_1)^2 + (u_2/b_2)^2 = \text{constant}.$$

The transition between the ellipses and the "plus signs" is, of course, continuous.

II.G. Distribution of U_1 when U_1 and U_2 are independent L-stable variables and $U_1 + U_2 = U$ is known

This conditional distribution can be obtained as the intersection between the surface that represents the joint density $p_0(u_1, u_2)$ and the plane $u_1 + u_2 = u$. Thus, the conditional distribution is unimodal for small u. For large u, it has two sharply distinct maxima located near $u_1 = 0$ and near $u_2 = 0$.

More precisely, the conditional density of U_1 is given by $p_1(u_1)p_2(u - u_1)/q(u)$, where $q(u)$ is the density of $U = U_1 + U_2$. Let u be positive and very large; if u_1 is small, one can use the scaling approximations for $p_2(u_2)$ and $q(u)$, obtaining

$$\frac{p_1(u_1)p_2(u-u_1)}{q(u)} \sim \frac{C'_1}{C'_1 + C'_2}\, p_1(u_1).$$

If u_2 is small, one similarly obtains

$$\frac{p_1(u_1)p_2(u-u_1)}{q(u)} \sim \frac{C'_2}{(C'_1 + C'_2)}\, p_2(u-u_1).$$

In other words, the conditional density $p_1(u_1)p_2(u - u_1)/q(u)$ looks as if two unconditioned distributions, scaled down in the ratios $C'_1/(C'_1 + C'_2)$ and $C'_2/(C'_1 + C'_2)$, had been placed near $u_1 = 0$ and $u_1 = u$. If u is negative, but very large in absolute value, a similar result holds with C''_1 and C''_2 replacing C'_1 and C'_2.

For example, for $\alpha = 2 - \varepsilon$ and $C'_1 = C'_2$, the conditional distribution is made up of two almost Gaussian bells, scaled down to one-half of their height. But, as α tends toward 2, these two bells become smaller and a third bell appears near $u_1 = u/2$. Ultimately, the two side bells vanish, leaving a single central bell. This limit corresponds to the fact that when the sum $U_1 + U_2$ is known, the conditional distribution of a Gaussian U_1 is itself Gaussian.

III.A. Explanation of large price changes as due to causal or random "contaminators"

One very common approach is to note that, a posteriori, large price changes are usually traceable to well-determined "causes," and should be eliminated before one attempts a stochastic model of the remainder. Such preliminary censorship obviously brings any distribution closer to the Gaussian. This is, for example, what happens when the study is limited to "quiet periods" of price change. Typically, however, no discontinuity is observed between the "outliers" and the rest of the distribution. In such cases, the notion of outlier is indeterminate and arbitrary. above censorship is therefore usually indeterminate.

Another popular and classical procedure assumes that observations are generated by a mixture of two normal distributions, one of which has a small weight but a large variance and is considered as a random "contaminator." In order to explain the sample behavior of the moments, it unfortunately becomes necessary to introduce a larger number of contaminators, and the simplicity of the model is destroyed.

III.B. Introduction of the scaling distribution to represent price changes

I propose to explain the erratic behavior of sample moments by assuming that the corresponding population moments are infinite. This is an approach that I used successfully in a number of other applications and which I explained and demonstrated in detail elsewhere.

In practice, the hypothesis that moment are infinite beyond some threshold value is hard to distinguish from the scaling distribution. Assume that the increment, for example,

$$L(t, 1) = \log_e Z(t+1) - \log_e Z(t)$$

is a random variable with infinite population moments beyond the first. This implies that $\int p(u)\ u^2 du$ diverges but $\int p(u)\ u du$ converges (the integrals being taken all the way to infinity). It is of course natural, at least in the first stage of heuristic motivating argument, to assume that $p(u)$ is somehow "well-behaved" for large u. If so, our two requirements mean that, as $u \to \infty$, $p(u)u^3$ tends to infinity and $p(u)u^2$ tends to zero.

In other words: $p(u)$ must somehow decrease faster than u^{-2} and slower than u^{-3}. The simplest analytical expressions of this type are asymptotically scaling. *This observation provided the first motivation of the* *present study*. It is surprising that I could find no record of earlier application of the scaling distribution to two-tailed phenomena.

My further motivation was more theoretical. Granted that the facts impose a revision of Bachelier's process, it would be simple indeed if one could at least preserve the following convenient feature of the Gaussian model. Let the increments,

$$L(t, T) = \log_e Z(t+T) - \log_e Z(t),$$

over days, weeks, months, and years. In the Gaussian case, they would have different scale parameters, but the same distribution. This distribution would also rule the fixed-base relatives. This naturally leads directly to the probabilists' concept of L-stability examined in Section II.

In other words, the facts concerning moments, together with a desire for a simple representation, led me to examine the logarithmic price relatives (for unsmoothed and unprocessed time series relative to very active speculative markets), and check whether or not they are L-stable. Cotton provided a good example, and the present paper will be limited to the examination of that case.

Additional studies. My theory also applies to many other commodities (such as wheat and other edible grains), to many securities (such as those of the railroads in their nineteenth-century heyday), and to interest rates such as those of call or time money. These examples were mentioned in my IBM Research Note NC-87 (dated March 26, 1962). Later papers {P.S. 1996: see M 1967j{E15}} shall discuss these examples, describe some properties of cotton prices that my model fails to predict correctly and deal with cases when few "outliers" are observed. It is natural in these cases to favor Bachelier's Gaussian model – a limiting case in my theory as well as its prototype.

III.C. Graphical method applied to cotton price changes

Let us first describe Figure 5. The horizontal scale u of lines 1*a*, 1*b*, and 1*c* is marked only on lower edge, and the horizontal scale u of lines 2*a*, 2*b*, and 2*c* is marked along the upper edge.

The vertical scale gives the following relative frequencies:

$$\text{(A)} \quad \begin{cases} (1a) & \text{Fr}\,\{\log_e Z(t+\text{ one day}) - \log_e Z(t) > u\}, \\ (2a) & \text{Fr}\,\{\log_e Z(t+\text{ one day}) - \log_e Z(t) < -u\}, \end{cases}$$

both for the daily closing prices of cotton in New York, 1900-1905. (Source: the United Stated Department of Agriculture.)

$$\text{(B)} \quad \begin{cases} (1b) & \mathrm{Fr}\ \{ \log_e Z(t + \text{ one day}) - \log_e Z(t) > u\}, \\ (2b) & \mathrm{Fr}\ \{ \log_e Z(t + \text{ one day}) - \log_e Z(t) < - u\}, \end{cases}$$

both for an index of daily closing prices of cotton in the United States, 1944-58. (Source: private communication from Hendrick S. Houthakker.)

$$\text{(C)} \quad \begin{cases} (1c) & \mathrm{Fr}\ \{ \log_e Z(t + \text{ one month}) - \log_e Z(t) > u\}, \\ (2c) & \mathrm{Fr}\ \{ \log_e Z(t + \text{ one month}) - \log_e Z(t) < - u\}, \end{cases}$$

both for the closing prices of cotton on the 15th of each month in New York, 1880-1940. (Source: private communication from the United States Department of Agriculture.)

The theoretical $\log \Pr\{U > u\}$, relative to $\delta = 0$, $\alpha = 1.7$, and $\beta = 0$, is plotted as a solid curve on the same graph for comparison.

If it were true that the various cotton prices are L-stable with $\delta = 0$, $\alpha = 1.7$ and $\beta = 0$, the various graphs should be horizontal translates of each other. To ascertain that, on cursory examination, the data are in close conformity with the predictions of my model, the reader is advised to proceed as follows: copy on a transparency the horizontal axis and the theoretical distribution and to move both horizontally until the theoretical curve is superimposed on one or another of the empirical graphs. The only discrepancy is observed for line 2*b*; it is slight and would imply an even greater departure from normality.

A closer examination reveals that the positive tails contain systematically fewer data than the negative tails, suggesting that β actually takes a small negative value. This is confirmed by the fact that the negative tails, but not the positive, begin by slightly "overshooting" their asymptote, creating the expected bulge.

III.D. Application of the graphical method to the study of changes in the distribution across time

Let us now look more closely at the labels of the various series examined in the previous section. Two of the graphs refer to daily changes of cotton prices, near 1900 and 1950, respectively. It is clear that these graphs do not coincide, but are horizontal translates of each other. This implies that

Moving on to Figure 7, we compare the distribution of the averages with that of actual monthly values. We see that, overall, they only differ by a horizontal translation to the left, as predicted in Section IIC. Actually, in order to apply the argument of that section, it would be necessary to rephrase it by replacing $Z(t)$ by $\log_e Z(t)$ throughout. However, the geometric and arithmetic averages of daily $Z(t)$ do not differ much in the case of medium-sized overall monthly changes of $Z(t)$.

But the largest changes between successive averages are smaller than predicted. This seems to suggest that the dependence between successive daily changes has less effect upon actual monthly changes than upon the regularity with which these changes are performed. {P.S. 1996: see Appendix I of this chapter.}

III.F. A new presentation of the evidence

I will now show that the evidence concerning daily changes of cotton price strengthens the evidence concerning monthly changes, and conversely.

The basic assumption of my argument is that successive daily changes of log (price) are independent. (This argument will thus have to be revised when the assumption is improved upon.) Moreover, the population second moment of $L(t)$ seems to be infinite, and the monthly or yearly price changes are patently nonGaussian. Hence, the problem of whether any limit theorem whatsoever applies to $\log_e Z(t+T) - \log_e Z(t)$ can also be answered *in theory* by examining whether the daily changes satisfy the Pareto-Doeblin-Gnedenko conditions. *In practice*, however, it is impossible to attain an infinitely large differencing interval T, or to ever verify any condition relative to an infinitely large value of the random variable u. Therefore, one must consider that a month or a year is infinitely long, and that the largest observed daily changes of $\log_e Z(t)$ are infinitely large. Under these circumstances, one can make the following inferences.

Inference from aggregation. The cotton price data concerning daily changes of $\log_e Z(t)$ appear to follow the weaker asymptotic? condition of Pareto-Doeblin-Gnedenko. Hence, from the property of L-stability, and according to Section IID, one should expect to find that, as T increases,

$$T^{-1/\alpha}\{\log_e Z(t+T) - \log_e Z(t) - T\,E[L(t,1)]\}$$

tends towards a L-stable variable with zero mean.

values of N are reached. Therefore, if one believes that the limit is rapidly attained, the functions $Q'(u)$ and $Q''(u)$ of daily changes must vary very little in the tails of the usual samples. In other words, it is necessary, after all, that daily price changes be asymptotically scaling.

Aggregation. Here, the difficulties are of a different order. From the mathematical viewpoint, the L-stable law should become increasingly accurate as T increases. Practically, however, there is no sense in even considering values of T as long as a century, because one cannot hope to get samples sufficiently long to have adequately inhabited tails. The year is an acceptable span for certain grains, but here the data present other problems. The long available yearly series do not consist of prices actually quoted on some market on a fixed day of each year, but are averages. These averages are based on small numbers of quotations, and are obtained by ill-known methods that are bound to have varied in time. From the viewpoint of economics, two much more fundamental difficulties arise for very large T. First of all, the model of independent daily L's eliminates from consideration every "trend," except perhaps the exponential growth or decay due to a nonvanishing δ. Many trends that are negligible on the daily basis would, however, be expected to be predominant on the monthly or yearly basis. For example, the effect of weather upon yearly changes of agricultural prices might be very different from the simple addition of speculative daily price movements.

The second difficulty lies in the "linear" character of the aggregation of successive L's used in my model. Since I use natural logarithms, a small $\log_e Z(t+T) - \log_e Z(t)$ will be indistinguishable from the relative price change $[Z(t+T) - Z(t)]/Z(t)$. The addition of small L's is therefore related to the so-called "principle of random proportionate effect." It also means that the stochastic mechanism of prices readjusts itself immediately to any level that $Z(t)$ may have attained. This assumption is quite usual, but very strong. In particular, I shall show that if one finds that $\log Z(t+ \text{ one week}) - \log Z(t)$ is very large, it is very likely that it differs little from the change relative to the single day of most rapid price variation (see Section VE); naturally, this conclusion only holds for independent L's. As a result, the greatest of N successive daily price changes will be so large that one may question both the use of $\log_e Z(t)$ and the independence of the L's.

There are other reasons (see Section IVB) to expect to find that a simple addition of speculative daily price changes predicts values too high for the price changes over periods such as whole months.

Given all these potential difficulties, I was frankly astonished by the quality of the prediction of my model concerning the distribution of the changes of cotton prices between the fifteenth of one month and the fifteenth of the next. The negative tail has the expected bulge, and even the most extreme changes of price can be extrapolated from the rest of the curve. Even the artificial excision of the Great Depression and similar periods would not affect the results very greatly.

It was therefore interesting to check whether the ratios between the scale coefficients, $C'(T)/C'(1)$ and $C''(T)/C''(1)$, were both equal to T, as predicted by my theory whenever the ratios of standard deviations $\sigma'(T)/\sigma'(s)$ and $\sigma''(T)/\sigma''(s)$ follow the $T^{1/\alpha}$ generalization of the "$T^{1/2}$ Law," which was referred to in Section IIB. If the ratios of the C parameters are different from T, their values may serve as a measure of the degree of dependence between successive $L(t, 1)$.

The above ratios were absurdly large in my original comparison between the daily changes near 1950 of the cotton prices collected by H. Houthakker, and the monthly changes between 1880 and 1940 of the prices given by the USDA. This suggested that the price varied less around 1950, when it was supported, than it had in earlier periods. Therefore, I also plotted the daily changes for the period near 1900, which was chosen haphazardly, but not actually at random. The new values of $C'(T)/C'(1)$ and $C''(T)/C''(1)$ became quite reasonable: they were equal to each other and to 18. In 1900, there were seven trading days per week, but they subsequently decreased to five. Besides, one cannot be too dogmatic about estimating $C'(T)/C'(1)$. Therefore, the behavior of this ratio indicated that the "apparent" number of trading days per month was somewhat smaller than the actual number.

{P.S. 1996. Actually, I had badly misread the data: cotton was *not* traded on Sundays in 1900, and correcting this error improved the fit of the M 1963 model; see Appendix IV to this Chapter.}

IV. WHY ONE SHOULD EXPECT TO FIND NONSENSE MOMENTS AND NONSENSE PERIODICITIES IN ECONOMIC TIME SERIES

IV.A. Behavior of second moments and failure of the least-squares method of forecasting

It is amusing to note that the first known nonGaussian stable law, namely, the Cauchy distribution, was introduced in the course of a study of the method of least squares. A surprisingly lively argument followed the reading of Cauchy 1853. In this argument, Bienaymé 1853 stressed that a method based upon the minimization of the sum of squares of sample deviations cannot reasonably be used if the expected value of this sum is known to be infinite. The same argument applies fully to the problem of least-squares smoothing of economic time series, when the "noise" follows a L-stable law other than that of Cauchy.

Similarly, consider the problem of least-squares forecasting, that is, of the minimization of the expected value of the square of the error of extrapolation. In the L-stable case, this expected value will be infinite for every forecast, so that the method is, at best, extremely questionable.

One can perhaps apply a method of "least ζ-power" of the forecasting error, where $\zeta < \alpha$, but such an approach would not have the formal simplicity of least squares manipulations. The most hopeful case is that of $\zeta = 1$, which corresponds to the minimization of the sum of absolute values of the errors of forecasting.

IV.B. Behavior of the sample kurtosis and its failure as a measure of the "peakedness" or "long-tailedness" of a distribution

Pearson proposed to measure the peakedness or long-tailedness of a distribution by the following quantity, call "kurtosis"

$$\text{kurtosis} = -3 + \frac{\text{fourth population moment}}{\text{square of the second population moment}} .$$

In the L-stable case with $0 < \alpha < 2$, the numerator and the denominator both have an infinite expected value. One can, however, show that the sample kurtosis + 3 behaves proportionately to the following "typical" value

$$\frac{\left(\frac{1}{N} \text{ (the most probable value of } \sum L^4\text{)}\right.}{\left\{\frac{1}{N} \text{(the most probable value of } \sum L^2\text{)}\right\}^2}$$
$$= \frac{(\text{a constant})N^{-1+4/\alpha}}{\{(\text{a constant})\ N^{-1+2/\alpha}\}^2} = (\text{a constant})N.$$

It follows that the kurtosis is expected to increase without bound as $N \to \infty$. For small N, things are less simple, but presumably quite similar.

In this light, examine Cootner 1962. This paper developed the tempting hypothesis that prices vary at random as long as they do not wander outside a "penumbra", defined as an interval that well-informed speculators view as reasonable. But random fluctuations triggered by ill-informed speculators will eventually let the price go too high or too low. When this happens, the operation of well-informed speculators will induce this price to come back within the "penumbra." If this view of the world were correct, one would conclude that the price changes over periods of, say, fourteen weeks would be smaller than expected if the contributing weekly changes were independent.

This theory is very attractive a priori, but could not be generally true because, in the case of cotton, it is not supported by the facts. As for Cootner's own justification, it is based upon the observation that the price changes of certain securities over periods of fourteen weeks have a much smaller kurtosis than one-week changes. Unfortunately, his sample contains 250-odd weekly changes and only 18 fourteen-week periods. Hence, on the basis of general evidence concerning speculative prices, I would have expected, a priori, to find a smaller kurtosis for the longer time increment. Also, Cootner's evidence is not a proof of his theory; other methods must be used in order to attack the still very open problem of the possible dependence between successive price changes.

IV.C. Method of spectral analysis of random time series

These days, applied mathematicians are frequently presented with the task of describing the stochastic mechanism capable of generating a given time series $u(t)$, known or presumed to be random. The first response to such a problem is usually to investigate what is obtained by applying a theory of the "second-order random process." That is, assuming that $E(U)=0$, one forms the sample covariance

$$r(\tau)=\frac{1}{N-\tau}\sum_{t=T^0+1}^{t=T^0+N-\tau}u(t)u(t+\tau),$$

which is used, somewhat indirectly, to evaluate the population covariance

$$R(\tau)=E\left[U(t)U(t+\tau)\right].$$

Of course, $R(\tau)$ is always assumed to be finite for all τ. The Fourier transform of $R(\tau)$ is the "spectral density" of the process $U(t)$, and rules the "harmonic decomposition" of $U(t)$ into a sum of sine and cosine terms.

Broadly speaking, this method has been very successful, though many small-sample problems remain unsolved. Its applications to economics have, however, been questionable even in the large-sample case. Within the context of my theory, there is, unfortunately, nothing surprising in this finding. Indeed,

$$2E[U(t)U(t+\tau)] = E[U(t)+U(t+\tau)]^2 - E[U(t)]^2 - E[U(t+\tau)]^2.$$

For time series covered by my model, the three variances on the right hand side are all infinite, so that spectral analysis loses its theoretical motivation. This is a fascinating problem, but I must postpone a more detailed examination of it.

V. SAMPLE FUNCTIONS GENERATED BY L-STABLE PROCESSES; SMALL-SAMPLE ESTIMATION OF THE MEAN "DRIFT"

The curves generated by L-stable processes present an even larger number of interesting formations than the curves generated by Bachelier's Brownian motion. If the price increase over a long period of time happens a posteriori to have been exceptionally large, one should expect, in a L-stable market, to find that most of this change occurred during only a few periods of especially high activity. That is, one will find in most cases that the majority of the contributing daily changes are distributed on a fairly symmetric curve, while a few especially high values fall way outside this curve. If the total increase is of the usual size, to the contrary, the daily changes will show no "outliers."

In this section these results will be used to solve one small-sample statistical problem, that of the estimation of the mean drift δ, when the other parameters are known. We shall see that there is no "sufficient statistic" for this problem, and that the maximum likelihood equation does not necessarily have a single root. This has severe consequences from the viewpoint of the very definition of the concept of "trend."

V.A. Some properties of sample paths of Brownian motion

The sample paths of Brownian motion very much "look like" the empirical curves of time variation of prices or of price indexes. This was noted by Bachelier and (independently of him and of each other) by several modern writers (see especially Working 1934, Kendall 1953, Osborne 1959 and Alexander 1964), At closer inspection, however, one sees very clearly the effect of the abnormal number of large positive and large negative changes of $\log_e Z(t)$. At still closer inspection, one finds that the differences concern some of the economically most interesting features of the generalized central-limit theorem of the calculus of probability. It is therefore necessary to discuss this question in detail, beginning with a review of some classical properties of Gaussian random variables.

Conditional distribution of a Gaussian addend $L(+\tau, 1)$, knowing the sum $L(t, T) = L(t, 1) + ... + L(t+T-1, 1)$. Let the probability density of $L(t, T)$ be

$$\frac{1}{\sqrt{2\pi\sigma^2 T}} \exp\left\{ -\frac{u - \delta T^2}{2T\sigma^2} \right\}.$$

It is then easy to see that, if one knows the value of u of $L(t, T)$, the density of any of the quantities $L(t+\tau, 1)$ is given by

$$\frac{1}{2\pi\sigma^2(T-1)/T} \exp\left\{ -\frac{(u' - u/T)^2}{2\sigma^2(T-1)/T} \right\}.$$

This means that each of the contributing $L(t+\tau, 1)$ equals u/T plus a Gaussian error term. For large T, that term has the same variance as the unconditioned $L(t, 1)$ – one can in fact prove that the value of u has little influence upon the size of the largest of those "noise terms." One can therefore say that, whatever its value, u is roughly uniformly distributed over the T time intervals, each contributing negligibly to the whole.

Sufficiency of u for the estimation of the mean drift δ from the $L(t+\tau, 1)$. In particular, δ has vanished from the distribution of any $L(t+\tau, 1)$ conditioned by the value of u. In the vocabulary of mathematical statistics u is a "sufficient statistic" for the estimation of δ from the values of all the $L(t+\tau, 1)$. That is, whichever method of estimation a statistician may favor, his estimate of δ must be a function of u alone. The knowledge of intermediate values of $\log_e Z(t+\tau)$ is of no help. Most methods recommend estimating δ from u/T and extrapolating the future linearly from the two known points, $\log_e Z(t)$ and $\log_e Z(t+T)$. Since the causes of any price movement can be traced backwards only if the movement is of sufficient size, all that one can explain in the Gaussian case is the mean drift interpreted as a trend. Bachelier's model, which assumes a zero mean for the price changes, can only represent the movement of prices once the broad causal parts or trends have been removed.

V.B. One value from a process of independent L-stable increments

Returning to the L-stable case, suppose that the values of γ, of β (or of C' and C'') and of α are known. The remaining parameter is the mean drift δ; one must estimate δ starting from the known $L(t, T) = \log_e Z(t+T) - \log_e Z(t)$.

The unbiased estimate of δ is $L(t, T)/T$, while the estimate matches the observed $L(t, T)$ to its a priori *most probable* value. The "bias" of the maximum likelihood is therefore given by an expression of the form $\gamma^{1/\alpha} f(\beta)$, where the function $f(\beta)$ must be determined from the numerical table of the L-stable densities. Since β is mostly manifested in the relative sizes of the tails, its evaluation requires very large samples, and the quality of predictions will depend greatly upon the quality of one's knowledge of the past.

It is, of course, not at all clear that anybody would wish the extrapolation to be unbiased with respect to the mean of the change of the *logarithm* of the price. Moreover, the bias of the maximum likelihood estimate comes principally from an underestimate of the size of changes that are so large as to be catastrophic. The forecaster may very well wish to treat such changes separately, and to take into account his private opinions about many things that are not included in the independent-increment model.

V.C. Two values from a L-stable process

Suppose now that T is even and that one knows $L(t, T/2)$ and $L(t+T/2, T/2)$, and thus also their sum $L(t, T)$. Section IIG has shown that when the value $u = L(t, T)$ is given, the conditional distribution of $L(t, T/2)$ depends very sharply upon u. This means that the total change u is not a sufficient statistic for the estimation of δ; in other words, the estimates of δ will be changed by the knowledge of $L(t, T/2)$ and $L(t+T/2, T/2)$.

Consider, for example, the most likely value δ. If $L(t, T/2)$ and $L(t+T/2, T/2)$ are of the same order of magnitude, this estimate will remain close to $L(t, T)/T$, as in the Gaussian case. But suppose that *the actually observed* values of $L(t, T/2)$ and $L(t + T/2, T/2)$ are very unequal, thus implying that at least one of these quantities is very different from their common mean and median. Such an event is most likely to occur when δ is close to the observed value of either $L(t + T/2, T/2)/(T/2)$ or $L(t, T/2)/(t/2)$.

As a result, the maximum likelihood equation for δ has two roots, one near $2L(t, T/2)/T$ and the other near $2L(t + T/2, T/2)/T$. That is, the maximum-likelihood procedure says that one of the available items of information should be neglected, since any weighted mean of the two recommended extrapolations is worse than either. But nothing says which item should be neglected.

It is clear that few economists will accept such advice. Some will stress that the most likely value of δ is actually nothing but the most probable value in the case of the uniform distribution of a priori probabilities of δ. But it seldom happens that a priori probabilities are uniformly distributed. It is also true, of course, that they are usually very poorly determined. In the present problem, however, the economist will not need to determine these a priori probabilities with any precision: it will be sufficient to choose the most likely *for him* of the two maximum-likelihood estimates.

An alternative approach (to be presented later in this paper) will argue that successive increments of $\log_e Z(t)$ are not really independent, so that the estimation of δ depends upon the order of the values of $L(t, T/2)$ and $L(t + T/2, T/2)$, as well as upon their sizes. This may help eliminate the indeterminacy of estimation.

A third alternative consists in abandoning the hypothesis that δ is the same for both changes $L(t, T/2)$ and $L(t + T/2, T/2)$. For example, if these changes are very unequal, one can fit the data better by assuming that the trend δ is not linear but parabolic. In a first approximation, extrapolation would then consist in choosing among the two maximum-likelihood estimates the one which is chronologically the latest. This is an example of a variety of configurations which would have been so unlikely in the Gaussian case that they would have been considered nonrandom, and would have been of help in extrapolation. In the L-stable case, however, their probability may be substantial.

V.D. Three values from the L-stable process

The number of possibilities increases rapidly with the sample size. Assume now that T is a multiple of 3, and consider $L(t, T/3)$, $L(t+T/3, T/3)$, and $L(t+2T/3, T/3)$. If these three quantities are of comparable size, the knowledge of $\log Z(t+T/3)$ and $\log Z(t+2T/3)$ will again bring little change to the estimate based upon $L(t, T)$.

But suppose that one datum is very large and the others are of much smaller and comparable sizes. Then, the likelihood will have two local maximums, well separated, but of sufficiently equal sizes as to make it impossible to dismiss the smaller one. The absolute maximum yields the estimate $\delta = (3/2T)$ (sum of the two small data); the smaller local maximum yields the estimate $\delta = (3/T)$ (the large datum).

Suppose, finally, that the three data are of very unequal sizes. Then the maximum likelihood equation has *three* roots.

This indeterminacy of maximum likelihood can again be lifted by one of the three methods of Section VC. For example, if only the middle datum is large, the methods of nonlinear extrapolation will suggest a logistic growth. If the data increase or decrease – when taken chronologically – a parabolic trend should be tried. Again, the probability of these configurations arising from chance under my model will be much greater than in the Gaussian case.

V.E. A large number of values from a L-stable process

Let us now jump to the case of a very large amount of data. In order to investigate the predictions of my L-stable model, we must first reexamine the meaning to be attached to the statement that, in order that a sum of random variables follow a central limit of probability, it is necessary that each of the addends be negligible relative to the sum.

It is quite true, of course, that one can speak of limit laws only if the value of the sum is not *dominated* by any single addend known in advance. That is, to study the limit of $A(N)\Sigma U_n - B(N)$, one must assume that, for every n, Pr $|A(N)U_n - B(N)/N| \geq \varepsilon$ tends to zero with $1/N$.

As each addend decreases with $1/N$, their number increases, however, and the condition of the preceding paragraph does not by itself insure that the largest of the $|A(N)U_n - B(N)/N|$ is negligible in comparison with the sum. As a matter of fact, the last condition is true only if the limit of the sum is Gaussian. In the scaling case, on the contrary, the ratios

$$\frac{\max |A(N)U_n - B(N)/N|}{A(N)\sum U_n - B(N)} \text{ and } \frac{\text{plex pssum of } k \text{ largest } |A(N)U_n - B(N)/N|}{A(N)\sum U_n - B(BN)}$$

tend to nonvanishing limits as N increases (Darling 1952 and Arov & Bobrov 1960). In particular, it can be proven that, when the sum $A(N)\Sigma U_n - B(N)$ happens to be large, the above ratios will be close to *one*.

Returning to a process with independent L-stable $L(t)$, we may say the following: If, knowing α, β, γ, and δ, one observes that $L(t, T =$ one month) is *not* large, the contribution of the day of largest price change is likely to be nonnegligible in relative value, but it will remain small in absolute value. For large but finite N, this will not differ too much from the Gaussian prediction that even the largest addend is negligible.

Suppose, however, that $L(t, T =$ one month) is *very* large. The scaling theory then predicts that the sum of few largest daily changes will be very close to the total $L(t, T)$. If one plots the frequencies of various values of $L(t, 1)$, conditioned by a known and very large value for $L(t, T)$, one should expect to find that the law of $L(t + \tau, 1)$ contains a few widely "outlying" values. However, if the outlying values are taken out, the conditioned distribution of $L(t + \tau, 1)$ should depend little upon the value of the conditioned $L(t, T)$. I believe this last prediction to be well satisfied by prices.

Implications concerning estimation. Suppose now that δ is unknown and that one has a large sample of $L(t + \tau, 1)$'s. The estimation procedure then consists of plotting the empirical histogram and translating it horizontally until its fit to the theoretical density curve has been optimized. One knows in advance that the best value will be very little influenced by the largest outliers. Hence, "rejection of the outliers" is fully justified in the present case, at least in its basic idea.

V.F. Conclusions concerning estimation

The observations made in the preceding sections seem to confirm some economists' feeling that prediction is feasible only if the sample size is both very large and stationary, or if the sample size is small but the sample values are of comparable sizes. One can also make predictions from a sample size of one, but here the availability of a unique estimator is due only to ignorance.

V.G. Causality and randomness in L-stable processes

We mentioned in Section VA that, in order to be "causally explainable," an economic change must be large enough to allow the economist to trace back the sequence of its causes. As a result, the only causal part of a Gaussian random function is the mean drift δ. The same is true of L-stable random functions when their changes happen to be roughly uniformly distributed.

But it is not true in the cases where $\log_e Z(t)$ varies greatly between the times t and $t+T$, changing mostly during a few of the contributing days. Then, the largest changes are sufficiently clear-cut, and are sufficiently separated from "noise," to be explained causally, just as well as the mean drift.

In other words, a careful observer of a L-stable random function will be able to extract causal parts from it. But if the total change of $\log_e Z(t)$ is neither very large nor very small, there will be a large degree of arbitrariness in this distinction between causal and random. Hence, it would not be possible to determine whether the predicted proportions of the two kinds of effects are empirically correct.

In sum, the distinction between the causal and the random areas is sharp in the Gaussian case and very diffuse in the L-stable case. This seems to me to be a strong recommendation in favor of the L-stable process as a model of speculative markets. Of course, I have not the slightest idea why the large price movements should be representable in this way by a simple extrapolation of movements of ordinary size. I have come to believe, however, that it is very desirable that both "trend" and "noise" be aspects of the same deeper "truth." At this point, we can adequately describe it but cannot provide an explanation. I am certainly not antagonistic to the goal of achieving a decomposition of economic "noise" into parts similar to the trend, and to link various series to each other. But, until we come close to this goal, we should be pleased to be able to represent some trends as similar to "noise."

V.H. Causality and randomness in aggregation "in parallel"

Borrowing a term from elementary electrical circuit theory, the addition of successive daily changes of a price may be denoted by the term "aggregation in series," the term "aggregation in parallel" applying to the operation

$$L(t, T) = \sum_{i=1}^{I} L(i, t, T), = \sum_{i=1}^{I} \sum_{\tau=0}^{T-1} L(i, t+\tau, 1),$$

where i refers to "events" that occur simultaneously during a given time interval such as T or 1.

In the Gaussian case, one should, of course, expect any occurrence of a large value for $L(t, T)$ to be traceable to a rare conjunction of large changes in all or most of the $L(i, t, T)$. In the L-stable case, one should, on the contrary, expect large changes $L(t, T)$ to be traceable to one or a small number, of the contributing $L(i, t, T)$. It seems obvious that the L-stable prediction is closer to the facts.

If we add up the two types of aggregation in a L-stable world, we see that a large $L(t, T)$ is likely to be traceable to the fact that $L(i, t+\tau, 1)$ happens to be very large for one or a few sets of values of i and of τ. These contributions would stand out sharply and be causally explainable. But after a while, they should rejoin the "noise" made up of the other factors. The next rapid change of $\log_e Z(t)$ should be due to other "causes." If a contribution is "trend-making," in the above sense, during a large number of time-increments, one will naturally doubt that it falls under the same theory as the fluctuations.

VI. PRICE VARIATIONS IN CONTINUOUS TIME AND THE THEORY OF SPECULATION

The main point of this section is to examine certain systems of speculation, which appear advantageous, and to show that, in fact, they cannot be followed in the case of price series generated by a L-stable process.

VI.A. Infinite divisibility of L-stable variables

In theory, it is possible to interpolate $L(t, 1)$ indefinitely. That is, for every N, one can consider that a L-stable increment

$$L(t, 1) = \log_e Z(t+1) - \log_e Z(t)$$

is the sum of N independent, identically distributed random variables. The only difference between those variables and $L(t, 1)$ is that the constants γ, C' and C'' are N times smaller in the parts than in the whole.

In fact, it is possible to interpolate the process of independent L-stable increments to continuous time, assuming that $L(t, dt)$ is a L-stable variable with a scale coefficient $\gamma(dt) = dt\,\gamma(1)$. This interpolated process is a very important "zeroth" order approximation to the actual price changes. That is, its predictions are without doubt modified by the mechanisms of the market, but they are very illuminating nonetheless.

VI.B. Path functions of a L-stable process in continuous time

Mathematical models of physical or of social sciences almost universally assume that all functions can safely be considered to be continuous and to have as many derivatives as one may wish. Contrary to this expectation, the functions generated by Bachelier have no derivatives, even though they are indeed continuous. In full mathematical rigor, "there is a probability equal to 1 that they are continuous but nondifferentiable almost everywhere, but price quotations are always rounded to simple fractions of the unit of currency. If only for this reason, we need not worry about mathematical rigor here.

In the scaling case things are quite different. If my process is interpolated to continuous t, the paths which it generates become discontinuous in every interval of time, however small (in full rigor, they become "almost surely almost everywhere discontinuous"). That is, most of their variation occurs through noninfinitesimal "jumps." Moreover, the number of jumps larger than u and located within a time increment T is given by the law $C'T\,|d(u^{-\alpha})|$.

Let us examine a few aspects of this discontinuity. Again, very small jumps of $\log_e Z(t)$ could not be perceived, since price quotations are always expressed in simple fractions. More interesting is the fact that there is a nonnegligible probability of witnessing a price jump so large that supply and demand cease to be matched. In other words, the L-stable model can be considered as predicting the occurrence of phenomena likely to force the market to close. In a Gaussian model, such large changes are so extremely unlikely that the occasional closure of the markets must be explained by nonstochastic considerations.

The most interesting fact is, however, the large probability predicted for medium-sized jumps by the L-stable model. Clearly, if those medium-sized movements were oscillatory, they could be eliminated by market mechanisms such as the activities of the specialists. But if the movement is all in one direction, market specialists could at best transform a discontinuity into a change that is rapid but progressive. On the other hand, very few transactions would then be expected at the intermediate smoothing prices. As a result, even if the price Z_0 is quoted transiently, it may be impossible to act rapidly enough to satisfy more than a minute fraction of orders to "sell at Z_0." In other words, a large number of intermediate prices are quoted even if $Z(t)$ performs a large jump in a short time; but they are likely to be so fleeting, and to apply to so few transactions, that they are irrelevant from the viewpoint of actually enforcing a "stop loss order" of any kind. In less extreme cases – as, for example, when borrowings are over-subscribed – the market may have to resort to special rules of allocation.

These remarks are the crux of my criticism of certain systematic trading methods: they would perhaps be very advantageous *if only they could* be followed systematically; but, in fact, they *cannot* be followed. I shall be content here with a discussion of one example of this kind of reasoning.

VI.C. The fairness of Alexander's "filter" game

Alexander 1964 has suggested the following rule of speculation: "If the market goes up 5%, go long and stay long until it moves down 5%, at which time sell and go short until it again goes up 5%."

This procedure is motivated by the fact that, according to Alexander's interpretation, data would suggest that "in speculative markets, price changes appear to follow a random walk over time; but ... if the market has moved up x%, it is likely to move up more than x% further before it moves down x%." He calls this phenomenon the "persistence of moves." Since there is no possible persistence of moves in any "random walk" with zero mean, we see that if Alexander's interpretation of facts were confirmed, it would force us to seek immediately a model better than the random walk.

In order to follow this rule, one must, of course, watch a price series continuously in time and buy and sell whenever its variation attains the prescribed value. In other words, this rule can be strictly followed if and only if the process $Z(t)$ generates continuous path functions, as for example in the original Gaussian process of Bachelier.

Alexander's procedure cannot be followed, however, in the case of my own first-approximation model of price change in which there is a probability equal to one that the first move *not smaller* than 5% is *greater* than 5% and *not equal* to 5%. It is therefore mandatory to modify the filter method: one can at best recommend buying or selling when moves of 5% are *first exceeded*. One can prove that the L-stable theory predicts that this is game also fair. Therefore, evidence – as interpreted by Alexander – would again suggest that one must go beyond the simple model of independent increments of price.

But Alexander's inference was actually based upon the discontinuous series constituted by the closing prices on successive days. He assumed that the intermediate prices could be interpolated by some continuous function of continuous time – the actual form of which need not be specified. That is, whenever there was a difference of *over* 5% between the closing price on day F' and day F'', Alexander implicitly assumed that there was at least one instance between these moments when the price had gone up *exactly* 5 per cent. He recommends buying at this instant, and he computes the empirical returns to the speculator as if he were able to follow this procedure.

For price series generated by my process, however, the price actually paid for a stock will almost always be *greater* than that corresponding to a 5% rise; hence the speculator will almost always have paid *more* than assumed in Alexander's evaluation of the returns. Similarly, the price received will almost always be *less* than suggested by Alexander. Hence, at best, Alexander overestimates the yield corresponding to his method of speculation and, at worst, the very impression that the yield is positive may be a delusion due to overoptimistic evaluation of what happens during the few most rapid price changes.

One can, of course, imagine contracts guaranteeing that the broker will charge (or credit) his client the actual price quotation nearest by excess (or default) to a price agreed upon, irrespective of whether the broker was able to perform the transaction at the price agreed upon. Such a system would make Alexander's procedure advantageous to the speculator, but the money he would be making, on the average, would come from his broker and not from the market, and brokerage fees would have to be such as to make the game at best fair in the long run.

VII. A MORE REFINED MODEL OF PRICE VARIATION, TAKING ACCOUNT OF SERIAL DEPENDENCE

Broadly speaking, the predictions of my main model seem to me to be reasonable. At closer inspection, however, one notes that large price changes are not isolated between periods of slow change; they rather tend to be the result of several fluctuations, some of which "overshoot" the final changes. Similarly, the movements of prices in periods of tranquility seem to be smoother than predicted by my process. In other words, large changes tend to be followed by large changes – of either sign – and small changes tend to be followed by small changes, so that the isolines of low probability of $[L(t, 1), L(t-1, 1)]$ are **X**-shaped. In the case of daily cotton prices, Hendrik S. Houthakker stressed this fact in several conferences and in private conversation.

Such an **X**-shape is easily obtained by a 90° rotation from the "+ shape" which was observed when $L(t, 1)$ and $L(t-1, 1)$ are statistically independent and symmetric (Figure 4). This rotation introduces the two expressions:

$$S(t) = (1/2)[L(t, 1) + L(t-1, 1)] = (1/2)[\log_e Z(t+1) - \log_e Z(t-1)]$$

and

$$D(t) = (1/2)[L(t, 1) - L(t-1, 1)]$$
$$= (1/2)[\log_e Z(t+1) - 2\log_e Z(t) + \log_e Z(t-1)].$$

It follows that in order to obtain **X**-shaped empirical isolines, it would be sufficient to assume that the first and second finite differences of $\log_e Z(t)$ are two L-stable random variables, independent of each other, and naturally of $\log_e Z(t)$ (Figure 4). Such a process is invariant by time inversion.

It is interesting to note that the distribution of $L(t, 1)$, conditioned by the known $L(t-1, 1)$, is asymptotically scaling with an exponent equal to $2\alpha + 1$. A derivation is given at the end of this section. For the cases, we are interested in, $\alpha > 1.5$, hence $2\alpha + 1 > 4$. It follows that the conditioned $L(t, 1)$ has a finite kurtosis; no L-stable law can be associated with it.

Let us then consider a Markovian process with the transition probability I have just introduced. If the initial $L(T^0, 1)$ is small, the first values of $L(t, 1)$ will be weakly asymptotic scaling with a high exponent $2\alpha + 1$, so that $\log_e Z(t)$ will begin by fluctuating much less rapidly than in the case of independent $L(t, 1)$. Eventually, however, a large $L(t^0, 1)$ will appear. Thereafter, $L(t, 1)$ will fluctuate for some time between values of the orders of magnitude of $L(t^0, 1)$ and $-L(t^0, 1)$. This will last long enough to compensate fully for the deficiency of large values during the period of slow variation. In other words, the occasional sharp changes of $L(t, 1)$ predicted by the model of independent $L(t, 1)$ are replaced by oscillatory periods, and the periods without sharp change are shown less fluctuating then when the $L(t, 1)$ are independent.

We see that, if α is to estimated correctly, periods of rapid changes of prices must be considered with the other periods. One *cannot* argue that they are "causally" explainable and ought to be eliminated before the "noise" is examined more closely. If one succeeded in eliminating all large changes in this way, one would indeed have a Gaussian-like remainder. But this remainder would be devoid of any significance.

Derivation of the value $2\alpha + 1$ *for the exponent*. Consider

$$\Pr\{L(t, 1) > u, \text{ when } w < L(t-1, 1) < w + dw\}.$$

This is the product by $(1/dw)$ of the integral of the probability density of $[L(t-1, 1)L(t, 1)]$, over a strip that differs infinitesimally from the zone defined by

$$S(t) > (u+w)/2; w + S(t) < D(t) < w + S(t) + dw.$$

Hence, if u is large as compared to w, the conditional probability in question is equal to the integral

$$\int_{(u+w)/2}^{\infty} C'\alpha s^{-(\alpha+1)} C'\alpha (s+w)^{-(\alpha+1)} ds \sim (2\alpha+1)^{-1}(C')^2 \alpha^2 2^{-(2\alpha+1)} u^{-(2\alpha+1)}.$$

&&&& POST-PUBLICATION APPENDICES &&&&

These four appendices from different sources serve different purposes.

APPENDIX I (1996): THE EFFECTS OF AVERAGING

The M 1963 model of price variation asserts that price changes between equally spaced *closing* times are L-stable random variables. As shown momentarily, the model also predicts that changes between monthly average prices are L-stable.

To the contrary, Figure 7 suggests that the tails are *shorter* than predicted and the text notes that this is a token of interdependence between successive price changes.

The incorrect prediction. If $L(0)=0$ and $L(t)$ has independent L-stable increments, consider the increment between the "future" average from 0 to t and the value at t. Integration by parts yields

$$\frac{1}{t}\int_0^t L(s)ds - L(t) = -\frac{1}{t}\int_0^t s dL(s).$$

The r.h.s. is a L-stable random variable for which (scale)$^\alpha$ equals

$$t^{-\alpha}\int_0^t s^\alpha ds = (\alpha+1)^{-1}t.$$

The "past increment" is independent of the "future increment," and follows the same distribution. So does the difference between the two

APPENDIX II (MOSTLY A QUOTE FROM FAMA & BLUME 1966): THE EXEMPLARY FALL OF ALEXANDER'S FILTER METHOD

Section VI C of M 1963b criticizes a rule of speculation suggested in Alexander 1961, but does not provide a revised analysis of Alexander's data. However, Alexander's filters did not survive this blow. The story was told by Fama and Blume 1966 in the following terms:

"Alexander's filter technique is a mechanical trading rule which attempts to apply more sophisticated criteria to identify movements in stock prices. An x% filter is defined as follows: If the daily closing price of a particular security moves up at least x per cent, buy and hold the security until its price moves down at least x% from a subsequent high, at which time simultaneously sell and go short. The short position is maintained until daily closing prices rises at least x% above a subsequent low at which time one covers and buys. Moves less than x% in either direction are ignored.

"Alexander formulated the filter technique to test the belief, widely held among market professionals, that prices adjust gradually to new information.

"The professional analysts operate in the belief that there exist certain trend generating facts, knowable today, that will guide a speculator to profit if only he can read them correctly. These facts are assumed to generate trends rather than instantaneous jumps because most of those trading in speculative markets have imperfect knowledge of these facts, and the future trend of price will result from a gradual spread of awareness of these facts throughout the market [Alexander 1961, p.7].

"For the filter technique, this means that for some values of x we would find that 'if the stock market has moved up x% it is likely to move up more than x per cent further before it moves down by x% ' [Alexander 1961, p.26].

"In his Table 7, Alexander 1961 reported tests of the filter technique for filters ranging in size from 5 to 50 per cent. The tests covered different time periods from 1897 to 1959 and involved closing "prices" for two indexes, the Dow-Jones Industrials from 1897 to 1929 and Standard and Poor's Industrials from 1929 to 1959. In general, filters of all different sizes and for all the different time periods yielded substantial profits – indeed profits significantly greater than those of the simple buy-and-hold policy. This led Alexander to conclude that the independence assumption of the random-walk model was not upheld by his data.

"M 1963b [Section VI.C] pointed out, however, that Alexander's computations incorporated biases which led to serious overstatement of the profitability of the filters. In each transaction Alexander assumed that this hypothetical trader could always buy at a prices exactly equal to the low plus x per cent and sell at the high minus x per cent. In fact, because of the frequency of large price jumps, the purchase price will often be somewhat higher than the low plus x per cent, while the sale price will often be below the high minus x per cent. The point is of central theoretical importance for the L-stable hypothesis.

"In his later paper [Alexander 1964, Table 1] Alexander reworked his earlier results to take account of this source of bias. In the corrected tests the profitability of the filter technique was drastically reduced.

"However, though his later work takes account of discontinuities in the price series, Alexander's results are still very difficult to interpret. The difficulties arise because it is impossible to adjust the commonly used price indexes for the effects of dividends. This will later be shown to introduce serious biases into filter results."

Fama & Blume 1966 applied Alexander's technique to series of daily closing prices for each individual security of the Dow-Jones Industrial Average. They concluded that the filter method *does not* work.

Thus, the filters are buried for good, but many "believers" never received this message.

APPENDIX III (1996): ESTIMATION BIAS AND OTHER REASONS FOR $\alpha > 2$

Chapter E10, reproducing M 1960i{E10}, is followed by Post-Publication Appendix IV, adapted from M 1963i{E10}. The body of the present chapter, M 1963b, was written near-simultaneously with that appendix, and very similar comments can be made here. That is, for α close to 2, the diagrams in Figure 3 are inverse S-shaped, therefore, easily mistaken for straight lines with a slope that is $> \alpha$, and even > 2.

A broader structure is presented in Chapters E1 and E6, within which α has no upper bound. Therefore, the remark in the preceding paragraph *must not* be misconstrued. Estimation bias is only one of several reasons why an empirical log log plot of price changes may have a slope that contradicts the restriction [1, 2] that is characteristic of L-stability with $EU < \infty$.

APPENDIX IV (M 1972b): CORRECTION OF AN ERROR IN VCSP

• ***Section foreward.*** The correction of an error in *VCSP* = M 1963b improved in the fit between the data and the M 1963 model, eliminating some pesky descrepancies that *VCSP* had pointed out as deserving a fresh look. •

Infinite variance and of non-Gaussian L-stable distribution of price differentials were introduced for the first time in M 1963b. The prime material on which both hypotheses were based came in part from H.S. Houthakker and in part from the United States Department of Agriculture; it concerned daily spot prices of cotton.

Since then, the usefulness of those hypotheses was confirmed by the study of many other records, both in my work and in that of others. But it has now come to my attention that part of my early evidence suffered from a serious error. In the data sheets received from the USDA, an important footnote had been trimmed off, and as a result they were misread. Numbers which I had interpreted as Sunday closing prices were actually weekly price averages. They were inserted in the blanks conveniently present in the data sheets. My admiring joke about hard-working American cotton dealers of 1900-1905 was backfiring; no one corrected me in public, but I shudder at some comments that must have been made in private about my credibility. The error affected part of Figure 5 of M 1963b: the curves 1*a* and 2*a* relative to that period were incorrect.

After several sleepless nights, this error was corrected, and the analysis was revised. I am happy to report that my conclusion was upheld, in fact, much simplified, and the fit between the theory and the data improved considerably. M 1963b{E14} noted numerous peculiarities that had led me to consider my hypotheses as no more than rough first approximations. For example, the simplest random-walk model implied that a monthly price change is the sum of independent daily price changes. In fact, as I was careful to note, such was the case only if one assumed that a month included an "apparent number of trading days ... smaller than the actual number." The theory also implied that, whenever a monthly price change is large, it is usually about equal to the largest contributing daily price change. In fact, instances when large monthly changes resulted from, say, three large daily changes (one up and two down, or conversely) were more numerous in the data than predicted. Both findings suggested that a strong negative dependence exists between successive price changes. Also, prices seemed to have been more volatile around 1900 than around 1950. After the data have been corrected, these peculiarities have disappeared. In particular, the corrected curves 1*a* and 2*a* are nearly indistinguishable from the corresponding curves 1*b* and 2*b* relative to the Houthhakker data concerning the period 1950-58.

APPENDIX V (M 1982c): A "CITATION CLASSIC"

• ***Section foreward.*** In 1982, the *Citation Index of the Institute of Scientific Information* determined that M 1963b had become a *Citation Classic*. *Current Contents/Social and Behavioral Sciences* invited me to comment, "emphasizing the human side of the research – how the project was initiated, any obstacles encountered, and why the work was highly cited." •

✦ **Abstract.** Changes of commodity and security prices are fitted excellently by the L-stable probability distributions. Their parameter α is the intrinsic measure of price volatility. The model also accounts for the amplitudes of major events in economic history. An unprecedented feature is that price changes have an infinite population variance. ✦

Early in 1961, coming to Harvard to give a seminar on my work on personal income distributions, I stepped into the office of my host, H. S. Houthakker. On his blackboard, I noticed a diagram that was nearly identical to one I was about to draw, but concerned a topic of which I knew nothing: the variability of the price of cotton. My host had given up his attempt to model this phenomenon and challenged me to take over.

In a few weeks, I had introduced a radically new approach. It preserved the random walk hypothesis that the market is like a lottery or a casino, with prices going up or down as if determined by the throw of dice. It also preserved the efficient marked hypothesis that the market's collective wisdom take account of all available information, hence, the price tomorrow and on any day thereafter will *on the average* equal today's price. The third basis of the usual model is that price changes follow the Gaussian distribution. All these hypotheses, due to Louis Bachelier 1900, were first taken seriously in 1960. The resulting theory, claiming that price (or its logarithm) follows a Brownian motion, would be mathematically convenient, but it fits the data badly.

Most importantly, the records of throws of a die appear unchanged statistically. In comparison, the records of competitive price changes "look *nonstationary*"; they involve countless configurations that seem too striking to be attributable to mere chance. A related observation: the histograms of price changes are very far from the Galton ogive; they are long-tailed to an astonishing degree, due to large excursions whose size is obviously of the highest interest.

My model replaces the customary Gaussian hypothesis with a more general one, while allowing the population variance of the price changes to be infinite. The model is time invariant, but it creates endless configurations, and accounts for all the data, including both the seemingly nonstationary features, and the seemingly nonrandom large excursions.

A visiting professorship of economics at Harvard, 1962-1963, was triggered by IBM Research Note NC-87 (M 1962i), which tackled the prices of cotton and diverse commodities and securities. Also, M 1963b was immediately reprinted in Cootner 1964, along with discussions by E. F. Fama, who was my student at the time, and by the editor. This publication must have affected my election to Fellowship in the Econometric Society. However, after a few further forays in economics, my interest was drawn irresistibly toward the very different task of creating a new fractal geometry of nature. Having learned to live with the unprecedented infinite variance syndrome had trained me to identify telltale signs of divergence in the most diverse contexts, and to account for them suitably.

By its style, my work on prices remains unique in economics; while all the other models borrow from the final formulas of physics, I lean on its basic mental tool (invariance principles) and deduce totally new formulas appropriate to the fact that prices are not subjected to inertia, hence need not be continuous. My work is also unique in its power: the huge bodies of data that it fits involve constant jumps and swings, but I manage to fit everything without postulating that the rules themselves shift and change. Thus, my models are acknowledged as having opened a path toward a new and more realistic economics. Nevertheless, the progress of this new economics is slow, due to inherent mathematical difficulties and to my failure to push its development. All too often (though with notable exceptions, e.g. in Samuelson 1967), a publication devoted to the totally discredited Gaussian model quotes my work largely to show the author's awareness of his own work's limitations, and possibly to assuage his conscience.

&&&&&&&&&&&& ANNOTATIONS &&&&&&&&&&&&

Comment on a generalized L-stability. To avoid unnecessary complication, this chapter does not write down the general definition of L-stability, which adds a constant to the right-hand side of (S). The effect is to introduce additional L-stable laws with $\alpha = 1$; they are of possible relevance in the study of business concentration (Chapter E13).

How this paper came to be written. This paper's publication story is unique in my record, since it led a serious economics journal to follow a practice that is reputed to characterize *The Reader's Digest*. The story begins with M 1962i, IBM Research Note NC-87, dated March 26, 1962. I followed that text when teaching in the Harvard Department of Economics in 1962-63 and it was discussed at the 1962 Annual Meeting of the Econometric Society (see Chapter E17).

Later in 1963, after I had moved from the Harvard Economics Department to the Harvard Applied Physics program, Paul Cootner called from MIT. His book of reprints on *The Random Character of Stock Market Prices* was nearing completion, and he was sorry I was not a contributor. He was very familiar with my Report NC-87 and also knew that I was planning to expand this report into a short book. He wanted me to contribute to his book a short text based on NC-87, but did not want to include unpublished pieces, only actual reprints. Could my work be rushed into print somewhere, anywhere, so he could then reprint it? His publisher would accept page proofs.

I called every economics journal editor in turn. Some asked me to spell my name, others inquired about my field of work, a few knew me, but bemoaned the length of their backlog and the slowness of their referees. Finally, I struck gold with Merton Miller, an editor of the *Journal of Business* of the University of Chicago. He asked for a few hours to check a few things, then called back with a deal: NC87 was already well-known in Chicago, therefore no refereeing was needed; if I could manage to mail a rough version of the paper within a week, he would stop an issue about to go to press, and add my paper to it. The journal would even provide editorial assistance, and there would be no bill for "excess" corrections in proof. This deal could not be turned down, and the paper and its reprint became widely known on the research side of the financial community. At one time, a reprint combining M 1963b and Fama 1963 was given as premium to new subscribers of the *Journal of Business*.

Belated acknowledgement. Only after the paper had appeared did Merton Miller tell me that the editor Miller selected was E. F. Fama, who was no longer my student but on the University of Chicago faculty. Had this information been available in advance, I would have acknowledged Fama's help in my paper. I thanked him verbally, but this was not enough. To thank him in writing, late is better than never.