The Variation of Certain Speculative Prices

Benoit Mandelbrot

The Journal of Business, Volume 36, Issue 4 (Oct., 1963), 394-419.

THE VARIATION OF CERTAIN SPECULATIVE PRICES*

BENOIT MANDELBROT†

I. INTRODUCTION

THE name of Louis Bachelier is often mentioned in books on diffusion process. Until very recently, however, few people realized that his early (1900) and path-breaking contribution was the construction of a random-walk model for security and commodity markets.[1] Bachelier's simplest and most important model goes as follows: let $Z(t)$ be the price of a stock, or of a unit of a commodity, at the end of time period t. Then it is assumed that successive differences of the form $Z(t+T) - Z(t)$ are independent, Gaussian or normally distributed, random variables with zero mean and variance proportional to the differencing interval T.[2]

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 because *the empirical distributions of price changes are usually too "peaked" to be relative to samples from Gaussian populations.*[3] That

* The theory developed in this paper is a natural continuation of my study of the distribution of income. I was still working on the latter when Hendrik S. Houthakker directed my interest toward the distribution of price changes. The present model was thus suggested by Houthakker's data; it was discussed with him all along and was first publicly presented at his seminar. I therefore owe him a great debt of gratitude.

The extensive computations required by this work were performed on the 7090 computer of the I.B.M. Research Center and were mostly programmed by N. J. Anthony, R. Coren, and F. L. Zarnfaller. Many of the data which I have used were most kindly supplied by F. Lowenstein and J. Donald of the Economic Statistics section of the United States Department of Agriculture. Some stages of the present work were supported in part by the Office of Naval Research, under contract number Nonr-3775(00), NR-047040.

† Harvard University and Research Center of the International Business Machines Corporation.

[1] The materials of this paper will be included in greater detail in my book tentatively titled *Studies in Speculation, Economics, and Statistics*, to be published within a year by John Wiley and Sons.

The present text is a modified version of my "Research Note," NC-87, issued on March 26, 1962 by the Research Center of the International Business Machines Corporation. I have been careful to avoid any change in substance, but certain parts of that exposition have been clarified, and I have omitted some less essential sections, paragraphs, or sentences. Sections I and II correspond roughly to chaps. i and ii of the original, Sections III and IV correspond to chaps. iv and v, Sections V and VI, to chap. vi, and Section VII, to chap. vii.

[2] The simple Bachelier 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, and for this reason many modern authors have suggested that the original assumption of independent increments of $Z(t)$ be replaced by the assumption of independent and Gaussian increments of $\log_e Z(t)$.

Since Bachelier's original work is fairly inaccessible, it is good to mention more than one reference: "Théorie de la spéculation" (Paris Doctoral Dissertation in Mathematics, March 29, 1900) *Annales de l'Ecole Normale Supérieure*, ser. 3, XVII (1900), 21–86; "Théorie mathématique du jeu," *Annales de l'Ecole Normale Supérieure*, ser. 3, XVIII (1901), 143–210; *Calcul des probabilités* (Paris: Gauthier-Villars, 1912); *Le jeu, la chance et le hasard* (Paris, 1914 [reprinted up to 1929 at least]).

[3] To the best of my knowledge, the first to note this fact was Wesley C. Mitchell, "The Making and Using of Index Numbers," Introduction to *Index Numbers and Wholesale Prices in the United States and Foreign Countries* (published in 1915 as Bulletin No. 173 of the U.S. Bureau of Labor Statistics, reprinted in 1921 as Bulletin No. 284). But unquestionable proof was only given by Maurice Olivier in "Les Nombres indices de la variation des prix" (Paris doctoral dissertation, 1926), and Frederick C. Mills in *The Behavior of Prices* (New York: National Bureau of Economic Research, 1927). Other evi-

is, the histograms of price changes are indeed unimodal and their central "bells" remind one 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, e.g., 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, and even though the series to which it applies is presumably stationary.

It is my opinion that these facts warrant a radically new approach to the problem of price variation.[4] The purpose of this paper will be to present and test such a new model of price behavior in speculative markets. The principal feature of this model is that starting from the Bachelier process as applied to $\log_e Z(t)$ instead of $Z(t)$, I shall replace the Gaussian distributions throughout by another family of probability laws, to be referred to as "stable Paretian," which were first described in Paul Lévy's classic *Calcul des probabilités* (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 stable Paretian 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 stable Paretian model will be examined. The remaining sections of the paper will then be devoted to a discussion of some of the more sophisticated mathematical and descriptive properties of the stable Paretian model. I shall, in particular, examine its bearing on the very possibility of implementing the stop-loss rules of speculation (Section VI).

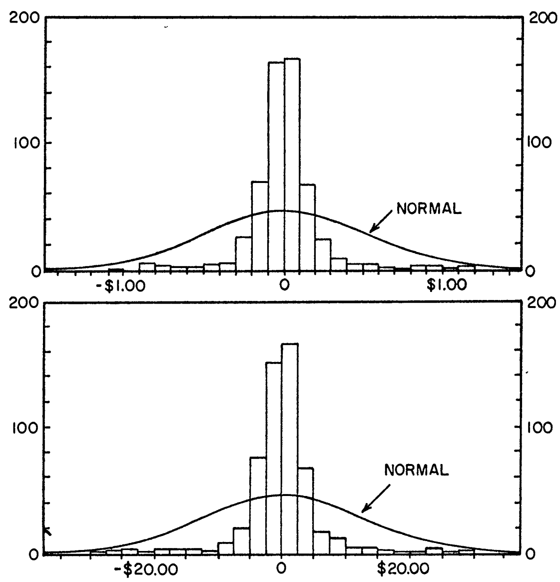

FIG. 1.—Two histograms illustrating departure from normality of the fifth and tenth difference of monthly wool prices, 1890–1937. In each case, the continuous bell-shaped curve represents the Gaussian "interpolate" based upon the sample variance. Source: Gerhard Tintner, *The Variate-Difference Method* (Bloomington, Ind., 1940).

dence, referring either to $Z(t)$ or to $\log_e Z(t)$ and plotted on various kinds of coordinates, can be found in Arnold Larson, "Measurement of a Random Process in Future Prices," *Food Research Institute Studies*, I (1960), 313–24; M. F. M. Osborne, "Brownian Motion in the Stock Market," *Operations Research*, VII (1959), 145–73, 807–11; S. S. Alexander, "Price Movements in Speculative Markets: Trends of Random Walks?" *Industrial Management Review of M.I.T.*, II, pt. 2 (1961), 7–26.

[4] Such an approach has also been necessary—and successful—in other contexts; for background information and many additional explanations see my "New Methods in Statistical Economics," *Journal of Political Economy*, Vol. LXXI (October, 1963).

I believe, however, that each of the applications should stand on their own feet and have minimized the number of cross references.

II. MATHEMATICAL TOOLS: PAUL LEVY'S STABLE PARETIAN LAWS

A. PROPERTY OF "STABILITY" OF THE GAUSSIAN LAW AND ITS GENERALIZATION

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 respectively equal to* σ'^2 *and* σ''^2. *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 mean square, is a solution to*

$$\text{(S)} \qquad s'U + s''U = sU \, ,$$

where s *is a function of* s′ *and* s″ *given by the auxiliary relation*

$$(\text{A}_2) \qquad s^2 = s'^2 + s''^2 \, .$$

It should be stressed that, from the viewpoint of 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 stability or invariance under addition, which is so fundamental in probability theory that it came to be referred to simply as "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

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

so that its density is of the form

$$dPr(U < u) = [\pi(1 + u^2)]^{-1} \, .$$

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

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

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

As to the general solution of equation

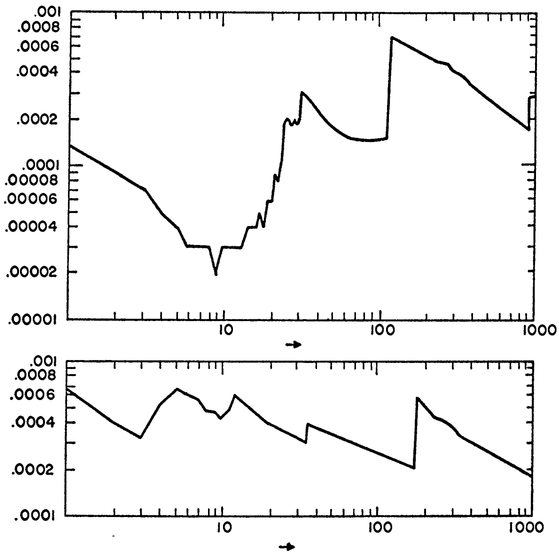

FIG. 2.—Both graphs are relative to the sequential sample second moment of cotton price changes. Horizontal scale represents time in days, with two different origins $T°$: on the upper graph, $T°$ was September 21, 1900; on the lower graph $T°$ was August 1, 1900. Vertical lines represent the value of the function

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

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, as privately reported by the United States Department of Agriculture.

(S), discovered by Paul Lévy,[5] the logarithm of its characteristic function takes the form

$$\text{(PL)} \log \int_{-\infty}^{\infty} \exp(iuz)\, dPr(U < u) = i\delta z - \gamma |z|^{\alpha}[1 + i\beta(z/|z|)\tan(\alpha\pi/2)],$$

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

Equation (PL) determines a family of distribution and density functions $Pr(U < u)$ and $dPr(U < u)$ that depend continuously upon four parameters which also happen to play the roles usually associated with the first four moments of U, as, for example, in Karl Pearson's classification.

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

One can say that α and β together determine the "type" of a stable random variable, and 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 having the same values for α, β and δ and having a value of γ equal to s^{α}: this means that the third parameter, γ, is a scale factor raised to the power α. Suppose now that U' and U'' are two independent stable variables, reduced and having the same values for α and β; since the characteristic function of $s'U' + s''U''$ is the product of those of $s'U'$ and of $s''U''$, the equation (S) is readily seen to be accompanied by the auxiliary relation

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

If on the contrary U' and U'' are stable with the same values of α, β and of $\delta = 0$, but with different values of γ (respectively, γ' and γ''), the sum $U' + U''$ is stable with the parameters α, β, $\gamma = \gamma' + \gamma''$ and $\delta = 0$. Thus the familiar additivity property of the Gaussian "variance" (defined by a mean-square) is now played by either γ or by a scale factor raised to the power α.

The final parameter of (PL) 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 δ has no obvious interpretation when $0 < \alpha < 1$ with $\beta \neq 0$.

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. Then, 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 it equals

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

so that 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 ran-

[5] Paul Lévy, *Calcul des probabilités* (Paris: Gauthier-Villars, 1925); Paul Lévy, *Théorie de l'addition des variables aléatoires* (Paris: Gauthier-Villars, 1937 [2d ed., 1954]). The most accessible source on these problems is, however, B. V. Gnedenko and A. N. Kolmogoroff, *Limit Distributions for Sums of Independent Random Variables*, trans. K. L. Chung (Reading, Mass.: Addison-Wesley Press, 1954).

dom variables. (This is, of 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 the change of $Z(t)$ over T days is equal to $\sigma(T) = T^{1/2}\sigma(1)$.

The corresponding prediction of my model is the following: 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 easy to find 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 $Z(t)$ over the time span $t^0 + 1$ to $t^0 + N$ can then be written as

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

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

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

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

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

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

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

As $N \to \infty$, one has

$$\gamma^0(N)/\gamma(U) \to 2N/(\alpha + 1) ,$$

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 $\alpha > 1$.

C. STABLE DISTRIBUTIONS AND THE LAW OF PARETO

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* non-Gaussian stable laws follow an asymptotic form of the law of Pareto, in the sense that there exist two constants $C' = \sigma'^\alpha$ and $C'' = \sigma''^\alpha$, linked by $\beta = (C' - C'')/(C' + C'')$, such that, when $u \to \infty$, $u^\alpha Pr(U > u) \to C' = \sigma'^\alpha$ and $u^\alpha Pr(U < -u) \to C'' = \sigma''^\alpha$.

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

In the extreme cases where $\beta = 1$ and hence $C'' = 0$ (respectively, where $\beta = -1$ and $C' = 0$), the negative tail (respectively, the positive tail) decreases faster than the law of Pareto of index α. In fact, one can prove[6] that it withers away even faster than the Gaussian density so that the extreme cases of stable laws are practically **J**-shaped. They play an important role in my theory of the distributions of personal income or of city sizes. A number of further properties of stable laws may therefore be found in my publications devoted to these topics.[7]

D. STABLE VARIABLES AS THE ONLY POSSIBLE LIMITS OF WEIGHTED SUMS OF INDEPENDENT IDENTICALLY DISTRIBUTED ADDENDS

The stability of the Gaussian law may be considered as being only a matter of convenience, and it is 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*

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

has a limit that is finite and is not reduced to a non-random 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 non-Gaussian, the linearly weighted sum

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

was seen to be *identical in law* to U_n, so that the "limit" of that expression is already attained for $N = 1$ and is a stable non-Gaussian law. Let us now suppose that U_n is asymptotically Paretian with $0 < \alpha < 2$, but not stable; one can show that the limit exists in a real sense, and that it is the stable Paretian law having the same value of α. Again 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. There is little sense in copying the readily available full mathematical arguments, and experience shows that what was intended to be an illuminating heuristic explanation often looks like another instance in which far-reaching conclusions are based

[6] A. V. Skorohod, "Asymptotic Formulas for Stable Distribution Laws," *Dokl. Ak. Nauk SSSR*, XCVIII (1954), 731–35, or *Select. Tranl. Math. Stat. Proba. Am. Math. Soc.* (1961), pp. 157–61.

[7] Benoit Mandelbrot, "The Pareto-Lévy Law and the Distribution of Income," *International Economic Review*, I (1960), 79–106, as amended in "The Stable Paretian Income Distribution, When the Apparent Exponent Is near Two," *International Economic Review*, IV (1963), 111–15; see also my "Stable Paretian Random Functions and the Multiplicative Variation of Income," *Econometrica*, XXIX (1961), 517–43, and "Paretian Distributions and Income Maximization," *Quarterly Journal of Economics*, LXXVI (1962), 57–85.

on loose thoughts. Let me therefore just quote the facts:

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 the asymptotic law of Pareto:[8]

The conditions of Pareto-Doeblin-Gnedenko.—Introduce the notations

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

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

The conditions of P-D-G require that (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 law of Pareto, 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 stability equation (S). More precisely, the limit is Gaussian if and only if the* U_n *has finite variance; the limit is stable non-Gaussian if and only if the conditions of Pareto-Doeblin-Gnedenko are satisfied for some* $0 < \alpha < 2$. *Then* $\beta = (C' - C'')/(C' + C'')$ *and* $A(N)$ *is determined by the requirement that*

$$N\,Pr[U > u\,A^{-1}(N)] \to C'u^{-\alpha} .$$

[8] See Gnedenko and Kolmogoroff, *op. cit.*, n. 4, p. 175, who use a notation that does not emphasize, as I hope to do, the relation between the law of Pareto and its present generalization.

(Whichever the value of α, the P-D-G 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 Paretian (but not stable) variables U' and U''. First of all, for $u > 0$, one can write

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

But it follows from the law of Pareto 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

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

$$P'' \sim C''^2 \alpha z \exp(-\alpha z) .$$

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 Pareto-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 stable Paretian limit with the exponent α and the skewness

$$\beta = (C'^2 + C''^2 - 2C'C'')/(C'^2 + C''^2 + 2C'C'') = [(C' - C'')/(C' + C'')]^2 \geq 0 .$$

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

E. SHAPE OF STABLE PARETIAN DISTRIBUTIONS OUTSIDE ASYMPTOTIC RANGE

The result of Section IIC should not hide the fact that the asymptotic behavior is seldom the main thing in the applications. For example, if the sample size is N, the orders of magnitude of the largest and smallest item are given by

$$N\,Pr[U > u^{+}(N)] = 1\,,$$

and

$$N\,Pr[U < -u^{-}(N)] = 1\,,$$

and the interesting values of u lie between $-u^{-}$ and u^{+}. Unfortunately, except in the cases of Gauss and of Cauchy and the case ($\alpha = \frac{1}{2}$; $\beta = 1$), there are no known closed expressions for the stable densities and the theory only says 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 most probable value and 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 law, whose density $[\pi(1 + u^2)]^{-1}$ is always *smaller* than the Paretian density $1/\pi u^2$ toward which it tends in relative value 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 apply 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 down very fast to negligible values. Hence, again by continuity, the graph for $\alpha = 2 - \epsilon$ must also begin by a rapid decrease. 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, say α^0, for which the doubly logarithmic graph begins by overshooting its asymptote. In the neighborhood of α^0, the asymptotic α can be measured as a slope even if the sample is small. If $\alpha < \alpha^0$, the asymptotic slope will be underestimated by the slope of small samples; for $\alpha > \alpha^0$ it will be overestimated. The numerical evaluation of the densities yields a value of α^0 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 its doubly logarithmic graph 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, α is between the critical values and the sample size is not large enough, the graphs of the two tails will have slightly different over-all apparent slopes.

F. JOINT DISTRIBUTION OF INDEPENDENT STABLE PARETIAN 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

$$p^0(u_1, u_2) = \alpha C_1' u_1^{-(\alpha+1)}\, \alpha C_2' u_2^{-(\alpha+1)} = \alpha^2 C_1' C_2' (u_1 u_2)^{-(\alpha+1)}\,.$$

Hence, the lines of equal probability are portions of the hyperbolas

$$u_1 u_2 = \text{constant}\,.$$

In the regions where either U_1 or U_2 is large (but not both), these bits of hyperbolas are linked together as in Figure 4. That is, the isolines of small probability have a characteristic "plus-sign" shape. On the contrary, when both U_1 and U_2 are small, $\log_e p_1(u_1)$ and $\log_e p_2(u_2)$ are near their maxima and therefore can be locally approximated by $a_1 - (u_1/b_1)^2$ and $a_2 - (u_2/b_2)^2$. Hence, the probability isolines are ellipses of the form

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

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

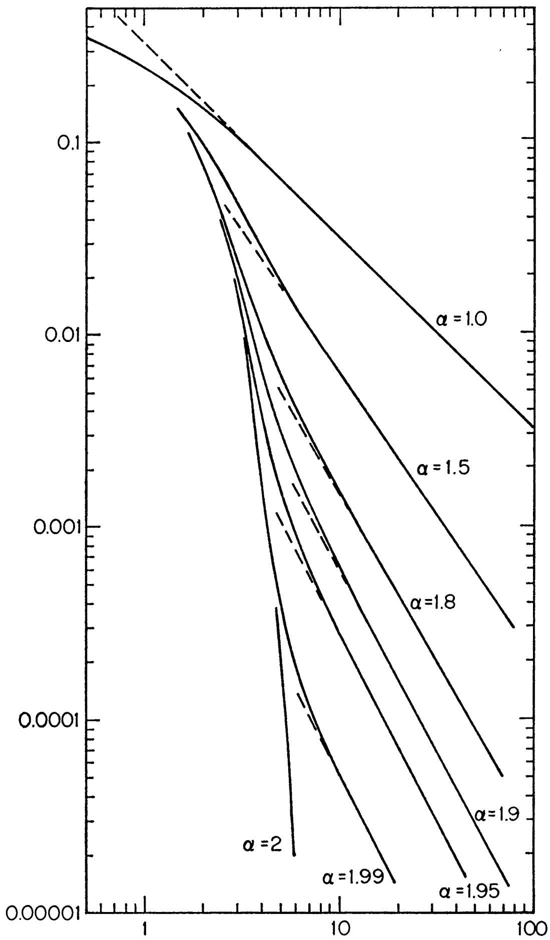

FIG. 3.—The various lines are doubly logarithmic plots of the symmetric stable Paretian probability distributions with $\delta = 0$, $\gamma = 1$, $\beta = 0$ and various values of α. Horizontally, $\log_e u$; vertically, $\log_e Pr(U > u) = \log_e Pr(U < -u)$. Sources: unpublished tables based upon numerical computations performed at the author's request by the I.B.M. Research Center.

G. DISTRIBUTION OF U_1, WHEN U_1 AND U_2 ARE INDEPENDENT STABLE PARETIAN 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 $u_1 + u_2 = u$. Hence 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 Paretian approximations for $p_2(u_2)$ and $q(u)$, obtaining

$$p_1(u_1)p_2(u - u_1)/q(u)$$
$$\sim [C_1'/(C_1' + C_2')]p_1(u_1) .$$

If u_2 is small, one similarly obtains

$$p_1(u_1)p_2(u - u_1)/q(u)$$
$$\sim [C_2'/C_1' + C_2')]p_2(u - u_1) .$$

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 $|u|$ is very large, a similar result holds with C_1'' and C_2'' replacing C_1' and C_2'.

For example, for $\alpha = 2 - \epsilon$ and $C_1' = C_2'$, the conditioned 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 and one is left with a central bell which 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. EMPIRICAL TESTS OF THE STABLE PARETIAN LAWS: COTTON PRICES

This section will have two main goals. First, from the viewpoint of statistical economics, its purpose is to motivate and develop a model of the variation of speculative prices based on the stable Paretian laws discussed in the previous section. Second, from the viewpoint of statistics considered as the theory of data analysis, I shall use the theorems concerning the sums ΣU_n to build a new test of the law of Pareto. Before moving on to the main points of the section, however, let us examine two alternative ways of treating the excessive numbers of large price changes usually observed in the data.

A. EXPLANATION OF LARGE PRICE CHANGES BY CAUSAL OR RANDOM "CONTAMINATORS"

One very common approach is to note that, a posteriori, large price changes are usually traceable to well-determined "causes" that 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 one restricts himself to the study of "quiet periods" of price change. There need not be any observable discontinuity between the "outliers" and the rest of the distribution, however, and the above censorship is therefore usually undeterminate.

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.

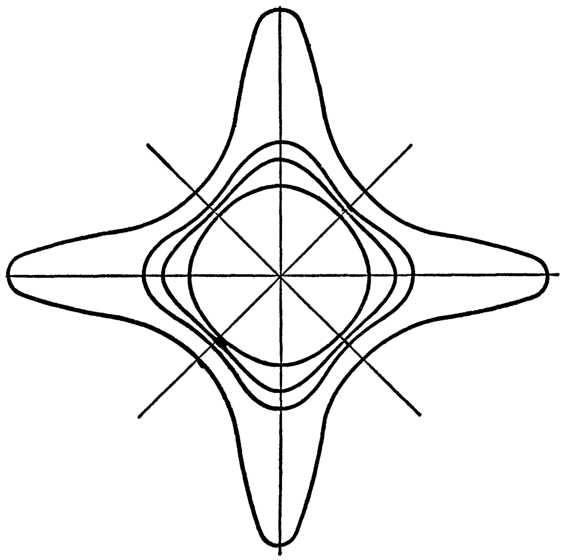

FIG. 4.—Joint distribution of successive price relatives $L(t, 1)$ and $L(t + 1, 1)$ under two alternative models. If $L(t, 1)$ and $L(t + 1, 1)$ are independent, they should be plotted along the horizontal and vertical coordinate axes. If $L(t, 1)$ and $L(t + 1, 1)$ are linked by the model in Section VII, they should be plotted along the bisectrixes, or else the above figure should be rotated by 45° before $L(t, 1)$ and $L(t + 1, 1)$ are plotted along the coordinate axes.

B. INTRODUCTION OF THE LAW OF PARETO TO REPRESENT PRICE CHANGES

I propose to explain the erratic behavior of sample moments by assuming that the population moments are infinite, an approach that I have used with success in a number of other applications and which I have explained and demonstrated in detail elsewhere.

This hypothesis amounts practically to the law of Pareto. Let us indeed assume that the increment

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

is a random variable with infinite population moments beyond the first. This implies that its density $p(u)$ is such 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 $u \rightarrow \infty$, $p(u)u^3$ tends to infinity and $p(u)u^2$ tends to zero.

In words: $p(u)$ must somehow decrease faster than u^{-2} and slower than u^{-3}. From the analytical viewpoint, the simplest expressions of this type are those with an asymptotically Paretian behavior. *This was the first motivation of the present study.* It is surprising that I could find no record of earlier application of the law of Pareto 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 convenient feature of the Gaussian model that the various increments,

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

depend upon T only to the extent of having different scale parameters. From all other viewpoints, price increments over days, weeks, months, and years would have the same distribution, which would also rule the fixed-base relatives. This naturally leads directly to the probabilists' concept of stability examined in Section II.

In other terms, the facts concerning moments, together with a desire to have a simple representation, suggested a check as to whether the logarithmic price relatives for unsmoothed and unprocessed time series relative to very active speculative markets are stable Paretian. Cotton provided a good example, and the present paper will be limited to the examination of that case. I have, however, also established that my theory 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.[9] On the other hand, there are unquestionably many economic phenomena for which much fewer "outliers" are observed, even though the available series are very long; it is natural in these cases to favor Bachelier's Gaussian model—known to be a limiting case in my theory as well as its prototype. I must, however, postpone a discussion of the limits of validity of my approach to the study of prices.

C. PARETO'S GRAPHICAL METHOD APPLIED TO COTTON-PRICE CHANGES

Let us begin by examining in Figure 5 the doubly logarithmic graphs of various kinds of cotton price changes as if they were independent of each other. The theoretical log $Pr(U > u)$, relative to $\delta = 0$, $\alpha = 1.7$, and $\beta = 0$, is plotted (*solid curve*) on the same graph for comparison. If the various cotton prices followed the stable Paretian law with $\delta = 0$, $\alpha = 1.7$ and $\beta = 0$, the various graphs should be horizontal translates of each other, and a cursory examination shows that the data are in close conformity with the predictions of my model. A closer examination suggests that the positive tails contain systematically fewer data than the negative tails, sug-

[9] These examples were mentioned in my 1962 "Research Note" (*op. cit.*, n. 1). My presentation, however, was too sketchy and could not be improved upon without modification of the substance of that "Note" as well as its form. I prefer to postpone examination of all the other examples as well as the search for the point at which my model of cotton prices ceases to predict the facts correctly. Both will be taken up in my forthcoming book (*op. cit.*, n. 1).

gesting that β actually takes a small negative value. This is also confirmed by the fact that the negative tails alone begin by slightly "overshooting" their asymptotes, creating the bulge that should be expected when α is greater than the critical value α° relative to one tail but not to the other.

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 near 1950, respectively. It is clear that these graphs do not coincide but are horizontal translates of each other. This implies that between 1900 and 1950 the generating process has changed only to the extent that its scale γ has become much smaller.

Our next test will be relative to monthly price changes over a longer time span. It would be best to examine the actual changes between, say, the middle of one

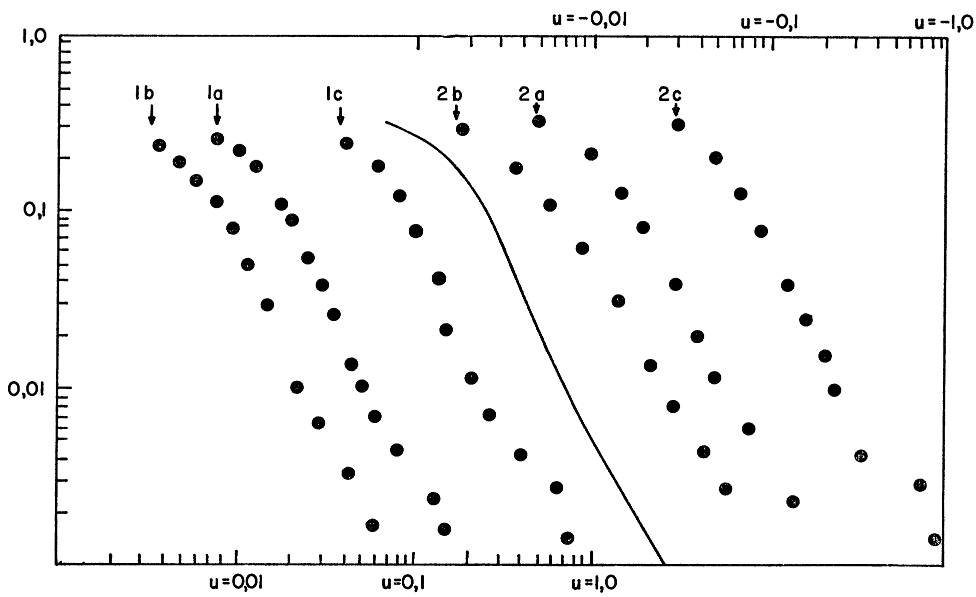

FIG. 5.—Composite of doubly logarithmic graphs of positive and negative tails for three kinds of cotton price relatives, together with cumulated density function of a stable distribution. Horizontal scale u of lines *1a*, *1b*, and *1c* is marked only on lower edge, and horizontal scale u of lines *2a*, *2b*, and *2c* is marked along upper edge. Vertical scale gives the following relative frequencies: (*1a*) $Fr[\log_e Z(t + \text{one day}) - \log_e Z(t) > u]$, (*2a*) $Fr[\log_e Z(t + \text{one day}) - \log_e Z(t) < -u]$, both for the daily closing prices of cotton in New York, 1900–1905 (source: private communication from the United States Department of Agriculture).

(*1b*) $Fr[\log_e Z(t + \text{one day}) - \log_e Z(t) > u]$, (*2b*) $Fr[\log_e Z(t + \text{one day}) - \log_e Z(t) < -u]$, both for an index of daily closing prices of cotton in the United States, 1944–58 (source: private communication from Hendrik S. Houthakker).

(*1c*) $Fr[\log_e Z(t + \text{one month}) - \log_e Z(t) > u]$, (*2c*) $Fr[\log_e Z(t + \text{one month}) - \log_e Z(t) < -u]$, 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 reader is advised to copy on a transparency the horizontal axis and the theoretical distribution and to move both horizontally until the theoretical curve is superimposed on either of the empirical graphs; the only discrepancy is observed for line *2b*; it is slight and would imply an even greater departure from normality.

month to the middle of the next. A longer sample is available, however, when one takes the reported monthly averages of the price of cotton; the graphs of Figure 6 were obtained in this way.

If cotton prices were indeed generated by a stationary stochastic process, our graphs should be straight, parallel, and uniformly spaced. However, each of the 15-year subsamples contains only 200-odd months, so that the separate graphs cannot be expected to be as straight as those relative to our usual samples of 1,000-odd items. The graphs of Figure 6 are, indeed, not quite as neat as those relating to longer periods; but, in the absence of accurate statistical tests, they seem adequately straight and uniformly spaced, except for the period 1880–96.

I conjecture therefore that, since 1816, the process generating cotton prices has changed only in its scale, with the possible exception of the Civil War and of the periods of controlled or supported prices. Long series of monthly price changes should therefore be represented by *mixtures* of stable Paretian laws; such mixtures remain Paretian.[10]

E. APPLICATION OF THE GRAPHICAL METHOD TO STUDY EFFECTS OF AVERAGING

It is, of course, possible to derive mathematically the expected distribution of the changes between successive monthly means of the highest and lowest quotation; but the result is so cumbersome as to be useless. I have, however, ascertained that the empirical distribution of these changes does not differ significantly from the distribution of the changes between the monthly means obtained by averaging all the daily closing quotations within months; one may therefore speak of a single average price for each month.

We then see on Figure 7 that the greater part of the distribution of the averages differs from that of actual monthly changes 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 over-all monthly changes of $Z(t)$).

However, 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.

F. A NEW PRESENTATION OF THE EVIDENCE

Let me now show that my evidence concerning daily changes of cotton price strengthens my 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 not Gaussian. 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 ever attain an infinitely large differencing interval T or to ever verify any condition relative to an infinitely large value of the random variable u. Hence one must consider that a month or a year is infinitely

[10] See my "New Methods in Statistical Economics," *Journal of Political Economy*, October, 1963.

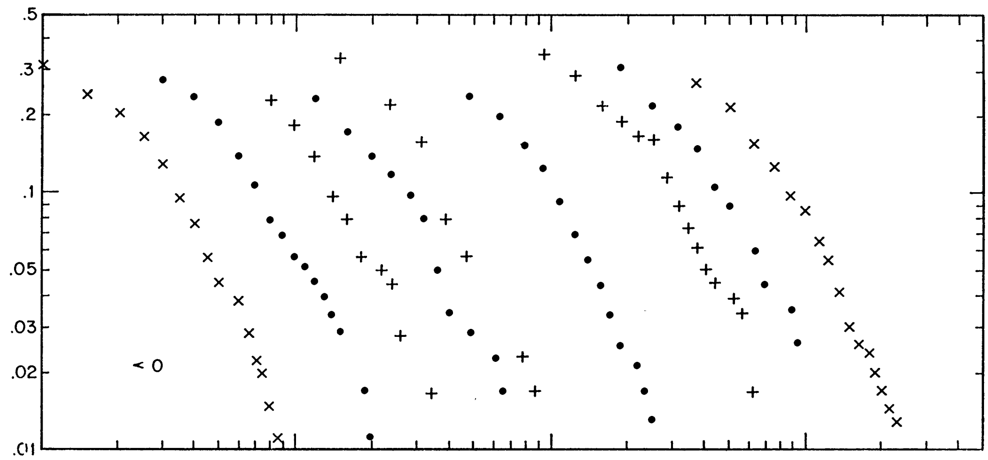

FIG. 6.—A rough test of stationarity for the process of change of cotton prices between 1816 and 1940. Horizontally, negative changes between successive monthly *averages* (source: *Statistical Bulletin* No. 99 of the Agricultural Economics Bureau, United States Department of Agriculture.) (To avoid interference between the various graphs, the horizontal scale of the kth graph from the left was multiplied by 2^{k-1}.) Vertically, relative frequencies $Fr(U < -u)$ corresponding respectively to the following periods (from left to right): 1816–60, 1816–32, 1832–47, 1847–61, 1880–96, 1896–1916, 1916–31, 1931–40, 1880–1940.

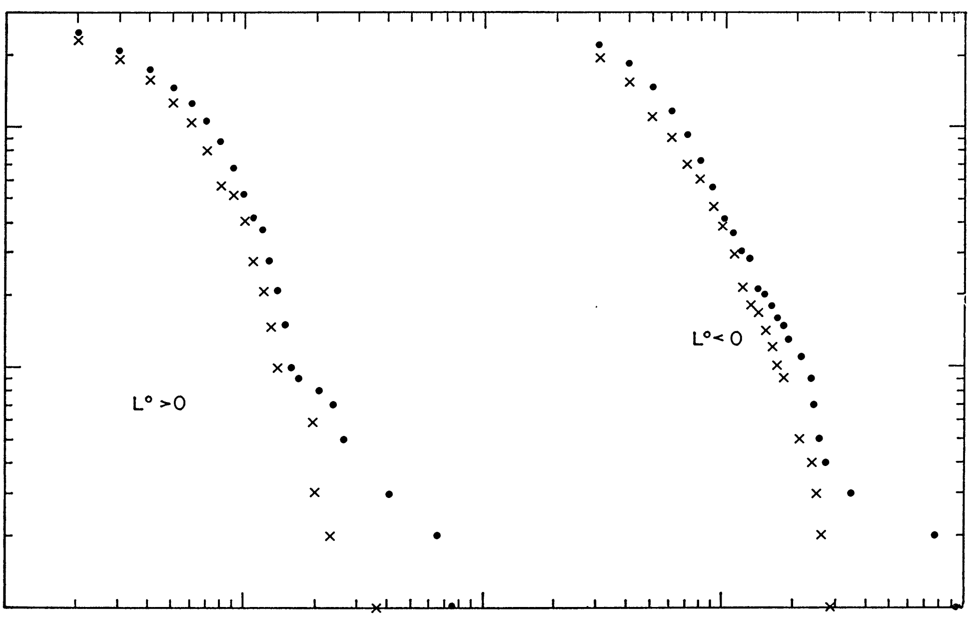

FIG. 7.—These graphs illustrate the effect of averaging. Dots reproduce the same data as the lines *1c* and *2c* of Fig. 5. The x's reproduce distribution of $\log_e Z^0(t+1) - \log_e Z^0(t)$, where $Z^0(t)$ is the average spot price of cotton in New York during the month t, as reported in the *Statistical Bulletin* No. 99 of the Agricultural Economics Bureau, United States Department of Agriculture.

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 $\log_e Z(t)$ surely appear to follow the weaker condition of Pareto-Doeblin-Gnedenko. Hence, from the property of 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 toward a stable Paretian variable with zero mean.

Inference from disaggregation.—Data seem to indicate that price changes over weeks and months follow the same law up to a change of scale. This law must therefore be one of the possible non-Gaussian limits, that is, it must be a stable Paretian. As a result, the inverse part of the theorem of Section IID shows that the daily changes of $\log Z(t)$ must satisfy the conditions of Pareto-Doeblin-Gnedenko.

It is pleasant to see that the inverse condition of P-D-G, which greatly embarrassed me in my work on the distribution of income, can be put to use in the theory of prices.

A few of the difficulties involved in making the above two inferences will now be discussed.

Disaggregation.—The P-D-G conditions are weaker than the asymptotic law of Pareto because they require that limits exist for $Q'(u)/Q''(u)$ and for $[Q'(u) + Q''(u)]/[Q'(ku) + Q''(ku)]$, but not for $Q'(u)$ and $Q''(u)$ taken separately. Suppose, however, that $Q'(u)$ and $Q''(u)$ still vary a great deal in the useful range of large daily variations of prices. If so, $A(N)\Sigma U_n - B(N)$ will not approach its own limit until *extremely* large 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 regions of the tails of the usual samples. In other words, it is necessary after all that the asymptotic law of Pareto apply to daily price changes.

Aggregation.—Here, the difficulties are of a different order. From the mathematical viewpoint, the stable Paretian 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 only if one is not worried by the fact that the long available series of yearly prices are ill known and variable averages of small numbers of quotations, not prices actually quoted on some market on a fixed day of each year.

From the viewpoint of economics, there are two much more fundamental difficulties with 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 non-vanishing 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, weather might have upon yearly changes of agricultural prices an effect 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 undistinguishable 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 predictions 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 are precise extrapolates from the rest of the curve. Even the artificial excision of the Great Depression and similar periods would not affect the general 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" referred to in Section IIB. If the ratios of the C parameter are different from T, their value 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 Houthakker and the monthly changes between 1880 and 1940 of the prices communicated by the USDA. This suggested that the supported prices around 1950 varied less than their earlier counterparts. Therefore I repeated the plot of daily changes for the period near 1900, chosen haphazardly but not actually at random. The new values of $C'(T)/C'(1)$ and $C''(T)/C''(1)$ became quite reasonable, equal to each other and to 18. In 1900, there were seven trading days per week, but they subsequently decreased to 5. 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.

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

A. BEHAVIOR OF SECOND MOMENTS AND FAILURE OF THE LEAST-SQUARES METHOD OF FORECASTING

It is amusing to note that the first known non-Gaussian stable law, namely, Cauchy's distribution, was introduced in the course of a study of the method of least squares. In a surprisingly lively argument following Cauchy's 1853 paper, J. Bienaymé[11] stressed that a method based upon the minimization of the sum

[11] J. Bienaymé, "Considérations a l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés," *Comptes rendus, Académie des Sciences de Paris*, XXXVII (August, 1853), 309–24 (esp. 321–23).

of squares of sample deviations cannot be reasonably 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 stable Paretian 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 stable Paretian 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.

B. BEHAVIOR OF THE KURTOSIS AND ITS FAILURE AS A MEASURE OF "PEAKEDNESS"

Pearson's index of "kurtosis" is defined as

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

If $0 < \alpha < 2$, the numerator and the denominator both have an infinite expected value. One can, however, show that the kurtosis behaves proportionately to its "typical" value given by

$$\frac{(1/N)\text{ (most probable value of } \Sigma L^4)}{[(1/N)\text{ (most probable value of } \Sigma L^2)]^2}$$

$$= \frac{\text{const. } N^{-1+4/\alpha}}{[\text{const. } N^{-1+2/\alpha}]^2} = \text{const. } N .$$

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

Let me examine the work of Cootner in this light.[12] He has developed the tempting hypothesis that prices vary at random only as long as they do not reach either an upper or a lower bound, that are considered by well-informed speculators to delimit an interval of reasonable values of the price. If and when ill-informed speculators let the price go too high or too low, the operations of the well-informed speculators will induce this price to come back within a "penumbra" à la Taussig. Under the circumstances, the price changes over periods of, say, fourteen weeks should be smaller than would be 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, and 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.

C. METHOD OF SPECTRAL ANALYSIS OF RANDOM TIME SERIES

Applied mathematicians are frequently presented these days with the task of describing the stochastic mecha-

[12] Paul H. Cootner, "Stock Prices: Random Walks vs. Finite Markov Chains," *Industrial Management Review of M.I.T.*, III (1962), 24–45.

nism capable of generating a given time series $u(t)$, known or presumed to be random. The response to such questions is usually to investigate first what is obtained by applying the theory of the "second-order random processes." That is, assuming that $E(U) = 0$, one forms the sample covariance

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

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

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

Of course, $R(\tau)$ is always assumed to be finite for all; its Fourier transform gives the "spectral density" of 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 such a finding. The expression $2E[U(t)U(t+\tau)]$ equals indeed $E[U(t) + U(t+\tau)]^2 - E[U(t)]^2 - E[U(t+\tau)]^2$; these three variances are all infinite for time series covered by my model, so that spectral analysis loses its theoretical motivation. I must, however, postpone a more detailed examination of this fascinating problem.

V. SAMPLE FUNCTIONS GENERATED BY STABLE PARETIAN PROCESSES; SMALL-SAMPLE ESTIMATION OF THE MEAN "DRIFT" OF SUCH A PROCESS

The curves generated by stable Paretian 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 usually large, in a stable Paretian market, one should expect to find that this change was mostly performed during 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 well outside this curve. If the total increase is of the usual size, the only difference will be that 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."

A. CERTAIN PROPERTIES OF SAMPLE PATHS OF BROWNIAN MOTION

As noted by Bachelier and (independently of him and of each other) by several modern writers,[13] the sample paths of the Brownian motion very much "look like" the empirical curves of time variation of prices or of price indexes. At closer inspection, however, one sees very well the effect of the abnormal number of

[13] See esp. Holbrolk Working, "A Random-Difference Series for Use in the Analysis of Time Series," *Journal of the American Statistical Association*, XXIX (1934), 11–24; Maurice Kendall, "The Analysis of Economic Time-Series—Part I: Prices," *Journal of the Royal Statistical Society*, Ser. A, CXVI (1953), 11–34; M. F. M. Osborne, "Brownian Motion in the Stock Market," *op. cit.*; Harry V. Roberts, "Stock-Market 'Patterns' and Financial Analysis: Methodological Suggestions," *Journal of Finance*, XIV (1959), 1–10; and S. S. Alexander, "Price Movements in Speculative Markets: Trends or Random Walks," *op. cit.*, n. 3.

large positive and 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 reminder of some classical facts concerning Gaussian random variables.

Conditional distribution of a Gaussian L(t), *knowing* L(t, T) = L(t, 1) + . . . + L(t + T − 1, 1).—Let the probability density of $L(t, T)$ be

$$(2\pi\sigma^2 T)^{-1/2} \exp[-(u - \delta T)^2/2T\sigma^2] .$$

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

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

We see 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, whichever 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 + τ, 1).—In particular, δ has vanished from the distribution of any $L(t + \tau, 1)$ conditioned by the value of u. This fact is expressed in mathematical statistics by saying that 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 to him. Most methods recommend estimating δ by 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 back only if it is ample enough, the only thing that can be explained in the Gaussian case is the mean drift interpreted as a trend, and 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.

B. SAMPLE FROM A PROCESS OF INDEPENDENT STABLE PARETIAN INCREMENTS

Returning to the stable Paretian case, suppose that one knows the values of γ and β (or of C' and C'') and of α. The remaining parameter is the mean drift δ, which one must estimate 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 maximum likelihood 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 tables of the stable Paretian densities. Since β is mostly manifested in the relative sizes of the tails, its evaluation requires very large samples, and the quality of one's 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 therefore very well wish to treat such changes separately and to take account of his private feelings about many things that are not included in the independent-increment model.

C. TWO SAMPLES FROM A STABLE PARETIAN PROCESS

Suppose now that T is even and that one knows $L(t, T/2)$ and $L(t+T/2, T/2)$ and their sum $L(t, T)$. We have seen in Section II G 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 either of $L(t+T/2, T/2)/(T/2)$ or of $L(t, T/2)/(T/2)$.

We see that as a result, the maximum likelihood equation for δ has two roots, respectively close to $2L(t, T/2)/T$ and to $2L(t+T/2, T/2)/T$. That is, the maximum-likelihood procedure says that one should neglect one of the available items of information, any weighted mean of the two recommended extrapolations being worse than either; but nothing says which item one should neglect.

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 a 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 may be tempted to believe that the trend δ is not linear but parabolic. Extrapolation would then approximately amount to 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 should be considered as non-random and would be of help in extrapolation. In the stable Paretian case, however, their probability may be substantial.

D. THREE SAMPLES FROM A STABLE PARETIAN 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 other are of much smaller and comparable sizes. Then, the likelihood equation will have two local maximums, having very different positions and sufficiently equal sizes 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 the middle datum only is large, the method of non-linear extrapolation will suggest a logistic growth. If the data increase or decrease—when taken chronologically—one will rather try a parabolic trend. Again the probability of these configurations arising from chance under my model will be much greater than in the Gaussian case.

E. A LARGE NUMBER OF SAMPLES FROM A STABLE PARETIAN PROCESS

Let us now jump to a very large number of data. In order to investigate the predictions of my stable Paretian model, we must first re-examine 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) $Pr|A(N)\ U_n - B(N)/N| > \epsilon)$ 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 their sum. As a matter of fact, the last condition is true only if the limit of the sum is Gaussian. In the Paretian case, on the contrary, the following ratios,

$$\frac{\max |A(N)\ U_n - B(N)/N|}{A(N)\Sigma U_n - B(N)}$$

and

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

tend to non-vanishing limits as N increases.[14] If one knows moreover that the sum $A(N)\Sigma U_n - B(N)$ happens to be large, one can prove that the above ratios should be expected to be close to *one.*

Returning to a process with independent stable Paretian $L(t)$, we may say the following: If, knowing α, β, γ, and δ, one observes that $L(t, T = \text{one month})$ is *not* large, the contribution of the day of largest price change is likely to be non-negligible 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 $L(t, T = \text{one month})$ is *very* large. The Paretian theory

[14] Donald Darling, "The Influence of the Maximum Term in the Addition of Independent Random Variables," *Transactions of the American Mathematical Society*, LXX (1952), 95–107; and D. Z. Arov and A. A. Bobrov, "The Extreme Members of Samples and Their Role in the Sum of Independent Variables," *Theory of Probability and Its Applications*, V (1960), 415–35.

then predicts that the sum of a 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 conditioning $L(t, T)$. I believe this last prediction to be very 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 consists in that case of plotting the empirical histogram and translating it horizontally until one has optimized its fit to the theoretical density curve. One knows in advance that this 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.

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 predict when the sample size is one, but here the unicity of the estimator is only due to ignorance.

G. CAUSALITY AND RANDOMNESS IN STABLE PARETIAN PROCESSES

We mentioned in Section V A that, in order to be "causally explainable," an economic change must at least 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. This will also apply to stable Paretian random functions when their changes happen to be roughly uniformly distributed.

Things are different when $\log_e Z(t)$ varies greatly between the times t an $t + T$, changing mostly during a few of the contributing days. Then, these largest changes are sufficiently clear-cut, and are sufficiently separated from "noise," to be traced back and explained causally, just as well as the mean drift.

In others words, a careful observer of a stable Paretian 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 one could not tell whether the predicted proportions of the two kinds of effects are empirically correct.

To sum up, the distinction between the causal and the random areas is sharp in the Gaussian case and very diffuse in the stable Paretian case. This seems to me to be a strong recommendation in favor of the stable Paretian process as a model of speculative markets. Of course, I have not the slightest idea why the large price movements should be represented in this way by a simple extrapolation of movements of ordinary size. I came to believe, however, that it is very desirable that both "trend" and "noise" be aspects of the same deeper "truth," which may not be explainable today, but which can be adequately described. I am surely not antagonistic to the ideal of economics: eventually to decompose even the "noise" into parts similar to the trend and to link various series to each other. But, until we can approximate this ideal, we can at least represent some trends as being similar to "noise."

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 designated 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 stable Paretian 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 Paretian prediction is closer to the facts.

To add up the two types of aggregation in a Paretian world, 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 of course rejoin the "noise" made up by 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, of course, doubt that it falls under the same theory as the fluctuations.

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

The main point of this section is to show that certain systems of speculation, which would have been advantageous if one could implement them, cannot in reality be followed in the case of price series generated by a Paretian process.

A. INFINITE DIVISIBILITY OF STABLE PARETIAN LAWS

Whichever N, it is possible to consider that a stable Paretian increment

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

is the sum of N independent, identically distributed, random variables, and that those variables differ from $L(t)$ only by the value of the constants γ, C' and C'', which are N times smaller.

In fact, it is possible to interpolate the process of independent stable Paretian increments to continuous time, assuming that $L(t, dt)$ is a stable Paretian variable with a scale coefficient $\gamma(dt) = dt\,\gamma(1)$. This interpolated process is a very important "*zero*th" order approximation to the actual price changes. That is, its predictions are surely modified by the mechanisms of the market, but they are very illuminating nonetheless.

B. PATH FUNCTIONS OF A STABLE PROCESS IN CONTINUOUS TIME

It is almost universally assumed, in mathematical models of physical or of social sciences, that all functions can safely be considered as being continuous and as having as many derivatives as one may wish. The functions generated by Bachelier are indeed continuous ("almost surely almost everywhere," but we may forget this qualification); although they have no derivatives ("almost surely almost nowhere"), we need not be concerned because price quotations are always rounded to simple fractions of the unit of currency.

In the Paretian case things are quite different. If my process is interpolated to continuous t, the paths which it generates become everywhere discontinuous (or rather, they become "almost surely almost everywhere discontinuous"). That is, most of their variation is performed through non-infinitesimal "jumps," the number of jumps larger than u and located within a time increment T, being 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. It is more interesting to note that there is a non-negligible probability that a jump of price is so large that "supply and demand" cease to be matched. In other words, the stable Paretian model may 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 non-stochastic considerations.

The most interesting fact is, however, the large probability predicted for medium-sized jumps by the stable Paretian 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 oversubscribed—the market may have to resort to special rules of allocation.

These remarks are the crux of my criticism of certain systematic methods: they would perhaps be very advantageous if only they could be enforced, but in fact they can only be enforced by very few traders. I shall be content here with a discussion of one example of this kind of reasoning.

C. THE FAIRNESS OF ALEXANDER'S GAME

S. S. Alexander 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%."[15]

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, one would have to look at a very early stage for a theoretical improvement over the random walk model.

In order to follow this rule, one must of course watch a price series continu-

[15] S. S. Alexander, *op. cit.* n. 3.

ously in time and buy or 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 per cent is *greater* than 5 per cent and *not equal* to 5 per cent. It is therefore mandatory to modify Alexander's scheme to suggest buying or selling when moves of 5 per cent are *first exceeded*. One can prove that the stable Paretian theory predicts that this game also is fair. Therefore, the 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 per cent 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 per cent rise; hence the speculator will almost always have paid more than assumed in Alexander's evaluation of the returns. On the contrary, 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

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 change. Similarly, the movement of prices in periods of tranquillity 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 private conversation.

Such an **X** shape can be easily obtained by rotation from the "plus-sign shape" which was observed in Figure 4 to be applicable when $L(t, 1)$ and $L(t-1, 1)$ are statistically independent and symmetric. The necessary 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 the **X** shape of the empirical isolines, it would be sufficient to assume that the first and second finite differences of $\log_e Z(t)$ are two stable Paretian random variables, independent of each other and naturally of $\log_e Z(t)$ (see Fig. 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 Paretian with an exponent equal to $2\alpha + 1$.[16] Since, for the usual range of α, $2\alpha + 1$ is greater than 4, it is clear that no stable Paretian law can be associated with the conditioned $L(t, 1)$. In fact, even the kurtosis is finite for the conditioned $L(t, 1)$.

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 Paretian 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 less fluctuating than when the $L(t, 1)$ are independent.

We see that, for the correct estimation of α, it is mandatory to avoid the elimination of periods of rapid change of prices. 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 have a Gaussian-like remainder which, however, would be devoid of any significance.

[16] Proof: $Pr[L(t, 1) > u$, when $w < L(t-1, 1) < w + dw]$ 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 following integral, carried from $(u + w)/2$ to ∞.

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