Journal of Business: 39, 1966, 242-255.

E19

Nonlinear forecasts, rational bubbles, and martingales

• ***Chapter foreword***. Two terms are found in the title of this reprint, but not of the originals, namely, "nonlinear" and "rational bubble." They express the two main points of this paper in words that were not available to me in 1966.

The main substantive finding was that rational behavior on the part of the market may lead to "wild" speculative bubbles of the form illustrated in Figure 1. The randomness of these bubbles is called "wild" in my present vocabulary, because they can be extremely large, and their sizes and duration follow a scaling distribution. This distribution is closely akin to the L-stable distribution introduced in the model of price variation presented in M 1963b{E14}.

Today, the existence of very large bubbles in actual market records is known to everyone: both to the economists and to the newspaper-reading public from New York to Tokyo. But it may be interesting that neither the term *bubble* nor the behavior it describes were known to me when I wrote this paper. Hence, my argument can be viewed as a *genuine prediction.*

A basic mechanism that is part of my argument was rediscovered in Blanchard & Watson 1982. But the bubbles examined by these authors are very different from those I predicted. They are not "wild" at all and can even be called "mild" because the distribution of size and duration is not scaling but exponential. (This is a possibility this paper considers briefly but then dismisses.) The duration of exponentially distributed bubbles would have a small scatter, and the largest among such bubbles would not be sufficiently distinct to be observed individually. In addition, Blanchard & Watson 1982 predicts that all bubbles end by a rise or a fall of the same size. This conclusion is closely akin to Bachelier's original idea that price changes are Gaussian, but is in disagreement with reality, which is altogether different.

Now we move on to the term "nonlinear" that was also added to the title. It does not concern the economics substance, but a mathematical technique. Wild bubbles appear when forecasting is nonlinear. Linear forecasting yields altogether different results investigated in M 1971e{E20}.

Given that both themes in this paper were ahead of their time, one hears with no surprise that it was originally viewed with trepidation and a certain fear. It was submitted to the *Journal of Political Economy*, but they pressed me to allow it to be transferred to the end of a special issue of the *Journal of Business*. The editor of that special issue, James H. Lorie, contributed *Comments* that remain interesting, and include statements that can serve as substitute to the *Abstract* that was lacking in my paper. •

✦ **In lieu of abstract: Comments by J. H. Lorie.** " [This] interesting article ... is almost purely theoretical and has no direct application to the selection of investments or the management of portfolios; however, it should prove to be very important. In the last few years a significant controversy has developed over whether the prices of stocks follow a "random walk." The proponents of this view – primarily academicians – have presented an impressive body of evidence, although it is by no means definitive. If they are to be believed, knowledge of the history of movements of the prices of stocks is of no value. Strongly opposing this

view are the technicians who believe that knowledge of the movement of stock prices, properly interpreted and usually considered in conjunction with information on volume, can yield extraordinarily high profit. The technicians have to believe that "tomorrow's" prices depend to some extent and in some way on "yesterday's" prices. There is some evidence of weak dependence. The random walkers say that it is too weak to be meaningful, while the technicians would assert the contrary ... Mandelbrot [shows] that it is theoretically possible for dependence to exist without knowledge of such dependence being valuable or capable of producing a profit.

"Of the work in progress on [diverse aspects of finance], much has had the effect of discrediting beliefs – and even some relatively sophisticated ones – about the behavior of security prices. Much of the work now in progress centers on the careful testing of more such beliefs, and I feel safe in predicting that the majority of the findings will be of the same general sort." ✦

THE BEHAVIOR OF SPECULATIVE PRICES has always been a subject of extreme interest. Most past work, including M 1963b{E14} emphasized the statistics of price changes. The present paper goes one step farther, and relates my earlier findings concerning the behavior of prices to more fundamental economic "triggering" quantities. This effort is founded on an examination, one that is simplified but demanding and detailed, of the roles that anticipation and expected utility play in economics.

1. INTRODUCTION AND SUMMARY OF EARLIER INVESTIGATIONS

My findings will depend on both the behavior of the underlying "triggering" variable and the relationship between the "triggering" variable and the price. It is possible to conceive of models where the price series follows a pure random walk, that is, price changes are *independent.* It is also possible to conceive of models where successive price changes are *dependent*. When prices do not follow a random walk, but the dependence cannot be used to increase expected profits, probabilists say that prices follow a "martingale process." Before exploring these intriguing possibilities, however, it is appropriate to begin with a brief review of the current state of affairs in the field.

When examining prices alone, one assumes implicitly that all other economic quantities are unknown and that their effects on the development of the price series $Z(t)$ are random. The stochastic mechanism that will generate the future values of $Z(t)$ may, however, depend on its past and present values. Insofar as the prices of securities or commodities are concerned, the strength of this dependence has long concerned market analysts and certain academic economists, and remarkably contradictory conclusions have evolved.

Among the market analysts, the technicians claim that a speculator can considerably improve his prospects of gain by correctly interpreting certain telltale "patterns" that a skilled eye can help him extract from the records of the past. This naturally implies that the future development of $Z(t)$ is greatly, although not exclusively, influenced by its past. It also implies that different traders, concentrating on different portions of the past record, should make different estimates of the future price $Z(t+T)$.

Academic economists tend to be skeptical of systematic trading schemes. An example is found in Section VI of M 1963b{E14} and its continuation in Fama & Blume 1966. These economists like to emphasize that, even if successive price changes were generated by tossing a fair coin, price series would include spurious "patterns." One should therefore expect that more elaborate probabilistic generating mechanisms could account for some other patterns as well and possibly even for all patterns. As a result, the basic attitude of economists is that the significance of any pattern must always be evaluated in the light of some stochastic model.

The earliest stochastic treatment of price behavior is found in the 1900 dissertation of Louis Bachelier. Bachelier 1900 conceived several models, of varying generality and complexity. His most general and least developed model states that the present price is an unbiased estimator of the price at any moment in the future. Bachelier's second-level model asserts that, for every t and T, the increment $Z(t+T)-Z(t)$ is independent of the values of Z up to and including time t. This assumption is best referred to as the "random walk." Bachelier's third-level model, the only one to be fully developed, asserts that $Z(t+T)-Z(t)$ is a Gaussian random variable with zero mean and a variance proportional to T. The present term for such a $Z(t)$ is "Brownian motion."

Despite its popularity, the Gaussian model clearly is contradicted by the facts. First of all, M 1963b{E14} focussed on the distribution of price changes and showed that price increments which are stable-Lévy accounts surprisingly well for many properties of extremely long price series. The present paper strives for an even better model, one generated by an explicit economic mechanism. The marginal distribution of price increments will be scaling, but the increments *will not* be independent. To be perfectly honest, an assumption of independence will creep in by the back door, through the hypotheses that will be made concerning the intervals between the instants when the weather changes. It would be easy to make less specific probabilistic assumptions but very hard to carry out their implications.

We shall find that the sample variation of price exhibits a variety of striking "patterns" but that these not benefit the trader, on the average.

The stochastic process $Z(t)$ to be examined will be a martingale. To define this concept, denote by t, $t+T$ and $\tilde{t}_i$ the present instant of time, a future instant and an arbitrary set of past instants. $Z(t)$ is a martingale if

$$E[Z(t+T), \text{ given the values of } Z(t) \text{ and of all the } Z(\tilde{t}_i)] = Z(t).$$

One immediate result of this definition is that $E[Z(t+T)$, given the value of $Z(t)] = Z(t)$. However, *much more* is implied in the martingale equality. It demands statistical independence between future anticipations and *all* past values of Z. Thus, one can define a martingale in two stages. (A) It is possible to speak of a single value for $E[Z(t+T)|Z(t)]$, without having to specify by which *past* values this expectation is conditioned. (B) One has $E[Z(t+T)|Z(t)] = Z(t)$. This two-stage definition should underline the central role that martingales are likely to play in the problem to which the present work is devoted: that of the usefulness of a knowledge of past prices for purposes of forecasting.

It should also be stressed that the *distribution* of $Z(t+T)$, conditioned by known values of $Z(t)$ and of the $Z(\tilde{t}_i)$, may very well depend upon the past values $Z(\tilde{t}_i)$. The expectation is unique in being unaffected by the $Z(\tilde{t}_i)$.

The application of martingales to price behavior gives meaning to the loose idea that prices are somehow "unbiased." This idea goes back at least to Bachelier, in whose mind "unbiasedness" meant that price determination in active speculative markets is governed by a linear utility function.

However, let us consider some nonlinear function F of the price. In general the expectation of $F[Z(t+T)]$ will *not* equal the present value $F[Z(t)]$. This means that, if our speculator's private utility function is not linear in Z, playing on Z may be advantageous or disadvantageous for him. Moreover, individual speculators need not be led by the same utility as the market considered as a whole. They may, for example, either seek or avoid a large dispersion of possible future prices $Z(t+T)$. Even in the case of a martingale, an increasingly detailed knowledge of the past may be useful for such purposes.

Similarly, if $\log Z$ is a martingale, playing on Z will be advantageous to speculators having a linear utility function. The fact that unbiasedness is linked to a choice of scale for Z is well known to mathematical statisticians.

Interest in martingales among pure probabilists is such that an immense variety of martingale processes has been described. If we dealt with a single economic series, namely the price, the choice among this wealth of possibility could only be directed by purely mathematical criteria – a notoriously poor guide. Hence, the present step beyond the random walk was undertaken only within the context of a "fundamental analysis," in which the price attempts to follow "value." That is, the present price $Z(t)$ is a function of past prices, and of the past and present values of the exogenous trigger $Y(t)$. In the present paper, the process generating value will be such that, as T increases, the expectation of $Y(t+T)$ will tend fairly rapidly toward a limit. Taking that limit to be the present price $Z(t)$, will achieve two results. (1) Price and value will occasionally coincide. (2) Price will be generated by a martingale stochastic model in which the present $Z(t)$ is an unbiased estimator of $Z(t+T)$. Moreover, for large enough values of T, $Z(t)$ is an unbiased estimator of $Y(t+T)$.

If, however, the process generating Y has other properties, the forecast future value $E[Y(t+\text{infinity})]$ need not be a martingale. An example to the contrary is given in Section IIG. Therefore, the fact that forecasting the value leads to a martingale in the price tells us something about the structure of the value as well as the structure of the market mechanism. If the forecasted value does not follow a martingale, price could follow a martingale only if they fail to follow forecasted value.

The above considerations are linked with the often-raised question of whether one can divide the speculators into several successive groups where members of the first group know only the present and past values of Z; members of the next group also know the present and past values of the single series Y, and know how the price will depend upon the variation of Y; members of further groups also know the temporal evolution of various series that contribute to Y, and again know how these series affect the price. In the model of the present paper knowing anything beyond the present $Z(t)$ brings no advantage, on the average.

Martingales are naturally closely related to other techniques of time-series analysis that involve conditional expectations, such as regression theory, correlation theory, and spectral representation. In particular, if $Z(t)$ is a martingale, its derivative is spectrally "white" in the sense that the covariance $C(\tau)$ between $Z'(t)$ and $Z'(t+\tau)$ vanishes if $\tau \neq 0$. It follows that the expected value of the sample spectral density of $Z'(t)$ will be a constant independent of frequency. A market that can associate such a series $Z(t)$ with the exogenous $Y(t)$, can be called a "whitener" of the derivative $Y'(t)$. However, one must keep in mind that spectral methods are concerned with measuring *correlation* rather than statistical *dependence*. Spectral whiteness expresses lack of correlation, but it is *not* synonymous with independence, except in one important but atypical case: when the marginal distribution and the joint distributions of prices at different times are Gaussian. Clearly, the examples I have constructed for this paper, are *not* Gaussian. In fact, whiteness is even weaker than the martingale identity.

II. THE FORECASTING FUNCTION OF EXCHANGE MARKETS AND THE PERSISTENCE OF PRICE MOVEMENT IN AGRICULTURE COMMODITIES

II.A. Statement of the problem

The present section will be devoted to the series of *equilibrium prices* for an agricultural commodity. Consideration of fluctuations around this series, due to temporal scatter of supply and demand, will be postponed until Section III. Here, the price $Z(t)$ will be equal to the expected value of the future crop, which in turn only depends upon past and future weather, according to the following five rules: (1) Weather can only be good, bad, or indifferent. (2) One is only interested in deviations of the price from some "norm," so that it is possible to neglect the price effects of indifferent weather. (3) When there were g good days and b bad days between moments t' and t'' within the growing season, the size of crop will have increased by an amount proportional to $g-b$. (4) Under the conditions of rule (3), the "value" $Y(t)$ of a unit quantity of the crop will have decreased by an amount proportional to $g-b$. (5) At any instant t, there is a single price of a unit quantity for future delivery, equal to

$$\lim_{T \to \infty} E[Y(t+T)].$$

Units will be so chosen that the price will increase by 1¢ when the ultimate expected value $Y(t)$ increases by the effect upon the crop of one day's bad weather. These rules are very simplified, and they do not even take into account the effect upon future prices of the portions of past crops that are kept in storage.

The total problem is so complex, however, that it is best to begin by following up each of its aspects separately.

It is readily acknowledged that the rules would be much more realistic if they referred to $\log Z$ instead of Z, and similarly to the logarithms of other quantities. This transformation was avoided, however, in order to avoid burdening the notation. The interested reader can easily make the transformation by himself.

Our rational forecast of Y naturally depends upon the weather forecast, i.e., upon the past of Y, the probability distribution of the lengths of the weather runs, and the rules of dependence between the lengths of successive runs.

The crudest assumption is to suppose that the lengths of the runs of good, bad, or indifferent weather are ruled by statistically independent exponential variables – as is the case if weather on successive days is determined by independent random events. Then the future discounted with knowledge of the past is exactly the same as the future discounted without knowledge of the past. In particular, if good and bad days are equally probable, the discounting of the future will not change the prices based upon the present crop size. This means that the process ruling the variation of $Z(t)$ is the simplest random walk, with equal probabilities for an increase or a decrease of price by 1¢.

Our "intuition" about the discounting of the future is of course based upon this case. But it is not necessary that the random variable U, designating the length of a good or bad run, be exponentially distributed. In all other cases, some degree of forecasting will be possible, so that the price will be influenced by the known structure of the process ruling the weather. The extent of this influence will depend upon the conditional distribution of the random variable U, when it is known that $U \geq h$. The following subsection will therefore discuss this problem.

II.B. The distribution of random variables conditioned by truncation

Exponential random variables. To begin with, let us note that the impossibility of forecasting in the exponential case can be restated as being an aspect of the following observation: Let U be the exponential random variable for which $P(u) = \Pr\{U \geq u\} = \exp(-bu)$, and let $U(h)$ designate the conditioned random variable U, conditioned by $U \geq h > 0$. The Bayes theorem, then, yields the following results: If $u < h$, one has $\Pr\{U(h) \geq u\} = 1$; if $u > h$, one has

$$\Pr\{U(h) \geq u\} = \Pr\{U \geq u \mid U \geq h\} = \frac{\exp(-bu)}{\exp(-bu)} = \exp[-b(u-h)].$$

This means that $U(h) - h$ is a random variable independent of h, but having a mean value $1/b$ determined by the original scale of the unconditioned U.

Uniformly scaling random variables. Assume now that the distribution of U is scaling. That is, two positive parameters σ and α are given. If $u < \sigma$, one has $\Pr\{U \geq u\} = 1$; if $u > \sigma$, one has $\Pr\{U \geq u\} = (u/\sigma)^{-\alpha}$. In the present case, Bayes's theorem yields the following results: If $h < \sigma$, one has $\Pr\{U(h) \geq u\} = \Pr\{U \geq u\}$; if $\sigma < u < h$, one has $\Pr\{U(h) \geq u\} = 1$; finally, if $\sigma < h < u$, one has

$$\Pr\{U(h) \geq u\} = \Pr\{U \geq u \mid U \geq h\} = \frac{(u/\sigma)^{-\alpha}}{(h/\sigma)^{-\alpha}} = (u/h)^{-\alpha}.$$

It is clear that the typical values of $U(h)$, such as the quantiles or the expectation, are proportional to h. For example, $hq^{-1/\alpha}$ gives the value of $U(h)$ that is exceeded with the probability q.

The mean of $U(h)$ is finite only if $\alpha > 1$. In that case, one has

$$E[U(h)] = \int_h^\infty \alpha h^\alpha u^{-\alpha} du = \frac{\alpha h}{(\alpha - 1)} .$$

$$E[U(h) - h] = \frac{h}{(\alpha - 1)} = \frac{E[U(h)]}{\alpha} .$$

This last quantity is smaller or greater than h according to whether α is smaller or greater than 2; if $\alpha = 2$, one finds $E[U(h) - h] = h$.

As to the marginal probability that $h < U < h + dh$ (knowing that $h < U$) it is equal to $\alpha h^{-(\alpha+1)} dh/h^{-\alpha} = \alpha dh/h$, which decreases with h.

In order to fully assess the above findings, it is helpful to contrast them with the result valid in the Gaussian case. As a simplified intermediate case, consider the random variable U for which

$\Pr\{U \geq u > 0\} = \exp(-bu^2)$. Then the arguments developed above show that, for $u > h$, one has

$$\Pr\{U(h) \geq u\} = \Pr\{U \geq u \mid U \geq h\} = \exp[-b(u^2 - h^2)] = \exp[-b(u+h)(u-h)].$$

It follows that all the typical values of $U(h) - h$, such as the expected value or the quantiles, are smaller than the mean, and are smaller than the quantiles of an auxiliary exponential variable $W^0(h)$ such that $\Pr\{W^0(h) \geq w\} = \exp(-2hbw)$. This shows that the mean of $U(h) - h$ is smaller than $1/2hb$, and therefore tends to zero as h tends to infinity.

Results are very similar in the Gaussian case, but the algebra is complicated and need not be given here.

An important property of the present conditioned or truncated variable $U(h)$ is that it is *scale-free* in the sense that its distribution does not depend upon the original scale factor σ. One may also say that the original scaling law is self-similar. Self-similarity is very systematically exploited in my studies of various empirical time series and spatial patterns. In particular, runs whose duration is scalingly distributed provide a very reasonable approximation to the "trend" component of a number of meteorological time series; and this is, of course, the motivation for their use in the present context.

Proof that the property of self-similarity uniquely characterizes the scaling distribution. Indeed, it means that the ratio $\Pr\{X \geq u\}/\Pr\{X \geq h\}$ be the same when X is the original variable U or the variable U divided by any positive number k. For this condition to be satisfied, the function $P(u) = \Pr\{U \geq u\}$ must satisfy $P(u)/P(h) = P(ku)/P(kh)$. Let $R = \log P$ be considered as a function of $v = \log u$. Then, the above requirement can then be written as

$$R(v) - R(v^0) = R(v + \log k) - R(v^0 + \log k).$$

This means that $R = \log P$ must be a linear function of $v = \log u$, which is, of course, the definition of the scaling law through doubly logarithmic paper, in the manner of Pareto.

II.C. Prices based upon a forecast crop size

Keeping the above preliminary in mind, let us return to the crop-forecasting problem raised in Section IIA, and assume that the lengths of successive weather runs are statistically independent random variables following the scaling distribution. It is clear that a knowledge of the past now becomes useful in predicting the future. The results become especially simple if one modifies the process slightly by assuming that weather alternates between "passive runs" or indifferent behavior, and "active runs" when it can be good or bad with equal probabilities. Then, as long as one is anywhere within a "passive run," prices will be unaffected by the number of indifferent days in the past. But if there have been h good or bad days in the immediate past, the same weather is likely to continue for a further period whose expected value is $h/(\alpha-1)$. (Things are actually slightly more complicated, as seen in the *Comment* at the end of this Subsection.)

Recall that the crop growth due to one day of good weather decreased the price by 1¢. A good day following h other good days will then decrease the price by the amount $[1+1/(\alpha-1)]¢=[\alpha/(\alpha-1)]¢$, in which the $1/(\alpha-1)¢$ portion is due to revised future prospects. But, when good weather finally turns to "indifferent," the price will go *up* by $h/(\alpha-1)$, to compensate for unfulfilled fears of future bounties. It should be noted that $h/(\alpha-1)$ *is not* a linear function of the known past values of Y. This implies that the best linear forecast in not optimal.

As a result, the record of the prices of our commodity will appear as a random alternation of three kinds of period, to be designated as "flat," "convex," and "concave," and defined as follows. During flat periods, prices will very little and "aimlessly." During concave periods, prices will go up by small equal amounts every day, yet, on the last day of the period they will fall by a fixed proportion of their total rise within the whole period. Precisely the opposite behavior will hold for convex periods.

Examples of these three kinds of periods have been shown in Figure 1. If a run of good weather is interrupted by a single indifferent day, the pattern of prices will be made up of a "slow fall, a rapid rise, a slow fall, and a rapid rise." Up to a small day's move, the point of arrival will be the same as if there had been no indifferent day in between; but that single day will "break" the expectations sufficiently to prevent prices from falling as low as they would have done in its absence.

Near the end of the growing season, the above forecasts should of course be modified to avoid discounting the weather beyond the harvest. If the necessary corrections are applied, the final price will precisely correspond to the crop size, determined by the difference between the number of days of good and bad weather. These corresponding corrections will not, however, be examined here.

Comment. Let us return to $h/(\alpha-1)$ for the expected value noted at the end of the first paragraph of this subsection. In fact a positive scaling random variable must have a minimum value $\sigma>0$; therefore, after an active run as started, its expected future length jumps to $\sigma/(\alpha-1)$ and stays there until the actual run length has exceeded σ. Such a fairly spurious jump will also appear in the exponential case if good weather could not follow bad weather, and conversely. One can, in fact, modify the process so as to eliminate this jump in all cases, but this would greatly complicate the formulas while providing little benefit.

It is also interesting to derive the forecast value of $E[Y(t+T)-Y(t)]$, when T is finite and the instant t is the hth instant of a bad weather run. One readily finds that

$$\frac{1}{h}E[Y(t+T)-Y(t)]=\alpha(\alpha-1)^{-1}[1-(1+T/h)^{1-\alpha}]-1.$$

This shows that the convergence of $E[Y(t+T)]$ to its asymptote is fast when h is small, and slow when h is large.

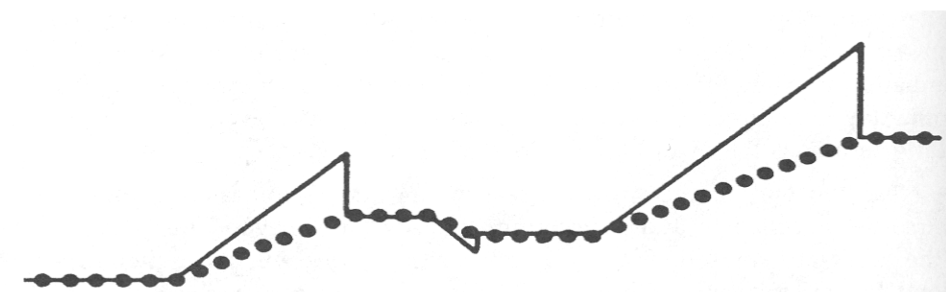

FIGURE C19-1. The abscissa is time for both the dotted and the bold lines; the ordinate is $Y(t)$ for the dotted line; The ordinate is $Z(t)$ for the bold line.

II.D. The martingale property of forecasted prices

The random series $Y(t)$ is *not* a martingale. To prove this fact, it suffices to exhibit one set of past values of Y for which the martingale property is not verified. We shall show that there is a nonvanishing conditioned expectation

$$E[Y(t+1) - Y(t), \text{ knowing the number } h \text{ of past good days}].$$

Proof. $Y(t+1) - Y(t) = 0$ if and only if the run of good weather breaks today. The probability of that event is $[h^{-\alpha} - (h+1)^{-\alpha}]/h^{-\alpha} \sim \alpha/h$. Otherwise, $Y(t+1) - Y(t) = -1$. Thus, the expectation of $Y(t+1) - Y(t)$ equals the probability that $Y(t+1) - Y(t) = -1$, which is $1 - \alpha/h$, a nonvanishing function of the past weather (whose history is fully represented for our present purpose by the duration of the current good weather run).

The price series $Z(t)$ *is a martingale.* To begin with, let us assume that h is known and evaluate the following conditioned increment:

$$-Z(t) + E[Z(t+1), \text{ given the value of } Z(t) \text{ and given that the number of preceding good days was exactly } h]$$

.

Let h be sufficiently large to avoid the difficulties due to the existence of a lower bound to the duration of a weather run. Two possibilities arise.

If weather continues to be good today, the price will go down by an amount equal to $\alpha/(\alpha-1)$. This event has a probability of $(h+1)^{-\alpha}/h^{-\alpha} \sim 1 - \alpha/h$.

If today's weather is indifferent, an event of probability α/h, the good-weather run is over and that the advance discounting of the effect of future weather was unwarranted. As a result, the price will climb up abruptly by an amount equal to $h/(\alpha-1)$.

The expected price change is thus approximately

$$\left(1 - \frac{\alpha}{h}\right) \frac{\alpha}{\alpha-1} - \frac{\alpha}{h} \frac{h}{\alpha-1} = \frac{1}{h} \alpha^2(\alpha-1),$$

which is approximately zero. The more involved rigorous derivation of the expected price change yields a value exactly (and not just approximately) equal to zero.

Now take into account the fact that one's actual knowledge of the past is usually not represented by the value of h but by some past values $Z(\tilde{t}_i)$

of $Z(t)$. The number of good days in the current run is then a random variable H, and $D(h) = \Pr\{H < h\}$ is a function determined by the values of the $Z(\tilde{t}_i)$. It follows that

$$E[Z(t+1), \text{given the value of } Z(t) \text{ and the past prices } Z(\tilde{t}_i)]$$
$$= \int dD(h) E[Z(t+1), \text{given the value of } Z(t) \text{ and the value of } h]$$
$$= \int dD(h) Z(t) = Z(t).$$

One shows very similarly that $E[Z(t+T)] = Z(t)$ when T exceeds 1, showing that $Z(t)$ *is indeed a martingale process*.

***Variance of* $\mathbf{Z(t+1) - Z(t)}$.** If the number of preceding good days was exactly h, this variance is equal to

$$\left(1 - \frac{\alpha}{h}\right)\left(\frac{\alpha}{\alpha - 1}\right)^2 + \frac{\alpha}{h}\left(\frac{h}{\alpha - 1}\right)^2,$$

which becomes proportional to h when h is large. Now suppose that h is not a known number, but is generated by a random variable H, that is conditioned by some known past prices $Z(\tilde{t}_i)$. Then the variance of $E(t+1) - Z(t)$ is proportional to $E(H)$.

If no past price is known, and $1 < \alpha < 2$, one can show that $E(H) = \infty$, and one falls back upon the infinite-variance property in M 1963b{E14}.

Comment. We have reached the climax of this story, and this is appropriate to comment again upon some observations made in the Introduction. If the price Z were generated by a random walk, then, whichever measure of risk has been adopted, no knowledge of the past should influence estimates of the risks involved in trading in Z. If, on the contrary, Z is generated by the present martingale, then the only "risk" that is not influenced by the past is constituted by the expectation of Z. A martingale is thus a "fair game." But, as h increases, so do the expected deviations from the expectation of $Z(t+1)$, and so do all other measures of "risk." This was to be expected, since, as h increases, so does the relative contribution to Z of anticipated changes in Y. Clearly, all traders, both risk-seekers and risk-avoiders, will want to know how the market value of a crop is divided between its present value and the changes anticipated before harvest time!

Also note the following: If the market is influenced more by risk-avoiders than by risk-seekers, the martingale equality should be replaced by $E[Z(t+T)] - Z(t) > 0$, the difference increasing with the variance of $Z(t+T)$. As a consequence, prices would increase in time, on the average, especially during the periods of high variance. However, this "tendency toward price increase" would be of significance only for traders who seek risk more than does the average trader on the market.

II.E. The distributions of price changes

This distribution is symmetric, thus it will suffice to derive it when $\Delta Z = Z(t+1) - Z(t)$ is positive or zero. We will denote by W' the mean duration of an indifferent weather run, and by W'' the mean duration of a good or bad weather run. Moreover, (for simplicity's sake) it will be assumed that W' and W'' are both large when measured in days.

The most significant price changes are those that satisfy $\Delta Z > \alpha/(\alpha - 1)$. These occur only on the last days of good weather runs, so that their total probability is $1/2(W' + W'')$. Their precise distribution is obtained by simply rescaling the law ruling the duration of good weather runs. Therefore, for $z > \alpha(\alpha - 1)$, (one has)

$$\Pr\{\Delta Z \geq z\} = \left\{ \frac{z+1}{\alpha/(\alpha-1)} \right\}^{-\alpha} \frac{1}{2(W' + W'')} .$$

Next, consider the probability that $Z' = 0$. This event occurs when t is anywhere within a run of indifferent weather, so that its probability is $W'/(W' + W'')$.

Finally, consider $\Delta Z = \alpha/(\alpha - 1)$ when the instant t is within a bad weather run but is not the last instant in that run. This event has the probability $(W'' - 1)/2(W' + W'')$.

The overall distribution of daily price changes is thus a "bell" with two scaling tails. It is shaped very much like a L-stable distribution, in this sense, the present model may be considered to provide a further elaboration of the process first proposed in M 1963b{E14}.

It is now safe to mention that the martingale property of forecast prices holds independently of the distribution $P(u)$ of bad weather runs, as long as runs are statistically independent. However, any non-scaling form of $P(u)$ would predict a marginal distribution of price change that is in conflict with the evidence brought forth in M 1963b{E14}.

II.F. A more involved agricultural commodity

Although still very crude, the preceding model seems more realistic than could have been expected. It can be further improved by taking into account the possibility of crop destruction by a natural calamity, such as hail. I have found that at least some natural calamities have scaling distributions. The extent of such calamities is presumably known only gradually, and they may therefore give rise to "patterns" similar to those we have studied above. The main interest of a mixture of several exogenous variables, however, is that it is unrealistic to believe that there is a proportionality between the distribution of large price changes and that of the time intervals between them. Such a proportionality holds in the case of a single trigger $Y(t)$, but not in the case of many triggers.

II.G. Best linear forecasts cannot be used to define prices

The following results, which I state without proof can be omitted without interrupting the continuity of the present work. Let us suppose that, instead of being ruled by the process $Y(t)$ that we have described, the value is ruled by a process $\tilde{Y}(t)$ with the following properties: $\Delta\tilde{Y}(t) = Y(t+1) - Y(t)$ is a stationary Gaussian process whose covariance function is equal to that of $\Delta Y(t)$. If so, the best extrapolate $\tilde{E}[\tilde{Y}(t+T)$, knowing $\tilde{Y}(t)$ for $s \leq t]$ is linear, and identical to the best linear extrapolate of $\tilde{Y}(t+T)$. As $T \to \infty$, this extrapolate tends to infinity therefore cannot define a price series $Z(t)$.

The above example suffices to show that, in order that the price based upon a forecast value *need not* be martingale. {P.S. 1996: This topic is discussed in M 1971e{E20}.}

III. PERSISTENCE OF PRICE MOVES FOR INDUSTRIAL SECURITIES

III.A. First approximation

The arguments of Section II can be directly translated into terms of "fundamental analysis" of security prices. Suppose, that the market value of a corporation is equal to the expected value of its future size X, computed while taking account of current and past values of its size $Y(t)$. If the rules of growth are of the form that we shall presently describe, it is meaningful to specify "the" expected future size by a single number, independent of the moment in the future to which one refers, and independent of the elements of the past history available for forecasting. The resulting theory is again greatly simplified (note the omission of all reference to current yield).

Our rules of growth are such that the lengths of periods of growth and decline are random, independent, and scaling. Thus, the longer a company has grown straight up, the more the outsiders should justifiably expect that it will grow in the future. Its market value Z should therefore justifiably increase by the multiple $1+1/(\alpha-1)$ of any additional growth actually observed for $Y(t)$. If, however, the growth of Y ever stops sufficiently for everyone to perceive it, one should observe a "break of confidence" and a fall of Z equal to the fraction $1/\alpha$ of the immediately preceding rise. If the growth of Y is stopped by "breathing spells," the growth of Z will have a sawtooth pattern. If a long growth period of Y, ending on a breathing spell, is modified by the addition of an intermediate breathing spell, the ultimate value of the company would be unchanged. But a single big tooth of Z would be replaced by two teeth, neither of which attains equally dizzying heights. In the absence of breathing spells, the price can go up and up, until the discounted future growth would have made the corporation bigger than the whole economy of its country, necessitating corrections that will not be examined in this paper.

Most of the further developments of this model would be very similar to those relative to the commodity examined in Section IIC. There is, however, a difference in that, if α is small, the expected length of the further growth period may be so long that one may need to discount the future growth at some nonvanishing rate.

III.B. Second approximation

Let us now examine the case of an industrial security whose fundamental value $X(t)$ follows a process of independent increments: either Bachelier's process of independent Gaussian increments, or the process in which the increments are L-stable (M 1963b{E14}. In both models, the rate of change of X may sometimes be very rapid; in the latter model it may even by instantaneous. But it will be assumed that the market only follows $X(t)$ through a smoothed-off form $Y(t)$ for which the maximum rate of change is fairly large, but finite. (In some cases, the establishment of an upper bound $\tilde{u}$ to the changes of Y may be the consequence of deliberate attempts to insure market continuity.)

In order to avoid mathematical complications, we will continue under the simplifying assumption that time is discrete. (The continuous time case is discussed at the end of this subsection.) In addition, assume that the maximum rate of change $\tilde{u}$ is known. It is clear that, whenever the market observes $Y(t) - Y(t-1) < \tilde{u}$, it will be certain that there was no smoothing off at time t and that $X(t) = Y(t)$. If $\tilde{u}$ is large enough, the equality $X = Y$ will hold for most values of t. Thus the market price $Z(t)$ will be equal most of the time to the fundamental value $X(t)$. Every so often, however, one will reach a point of time where $Y(t) - Y(t-1) = \tilde{u}$, a circumstance that may be due to any change $X(t) - X(t-1) \geq \tilde{u}$. At such instants, the value of $X(t) - X(t-1)$ will be greater than the observed value of $Y(t) - Y(t-1)$, and its conditional distribution will be given by the arguments of Section IIB; it will therefore critically depend upon the distribution of $X(t) - X(t-1)$.

If the distribution of the increments of X is Gaussian, and $\tilde{u}$ is large, the distributions of $X(t) - X(t-1)$, assuming that it is at least equal to $\tilde{u}$, will be clustered near $\tilde{u}$ as will the distribution of $X(t+1) - X(t-1)$. Hence, there will be a probability very close to 1 that $X(t+1) - X(t-1)$ will be smaller than $2\tilde{u}$ and $X(t+1) - X(t)$ will be smaller than $\tilde{u}$. As a result, $Y(t+1)$, will equal $X(t+1)$ and $Z(t+1)$ will be matched to $X(t+1)$. In other words, the mismatch between Z and Y will be small and short-lived in most cases.

Suppose now that the distribution of ΔX has two scaling tails with $\alpha < 2$. If $Y(t) - Y(t-1) = \tilde{u}$, while $Y(t-1) - Y(t-2) < \tilde{u}$, one knows that $X(t) - X(t-1) > \tilde{u}$ has a conditional expectation that is independent of the scale of the original process equal to $\alpha\tilde{u}/(\alpha - 1)$. The market price increment $Z(t) - Z(t-1)$ should therefore amplify, by the factor $\alpha/(\alpha - 1)$, the increment $Y(t) - Y(t-1)$ of the smoothed-off fundamental value.

Now, proceed to time $t+1$ and distinguish two cases: If $Y(t+1) - Y(t) < \tilde{u}$, the market will know that $X(t+1) = Y(t+1)$. Then the price $Z(t+1)$ will equal $X(t+1) = Y(t+1)$, thus falling from the inflated anticipation equal to $X(t-1) + \alpha\tilde{u}/(\alpha - 1)$. But if $Y(t+1) - Y(t) = \tilde{u}$, the market will know that $X(t+1) - X(t-1) = Y(t-1) + 2\tilde{u}$.

It follows that the conditioned difference $X(t) - X(t-1)$ will be close to following a scaling distribution truncated to values greater than $2\tilde{u}$. Thus the expected value of $X(t+1)$ – which is also the market price $Z(t+1)$ – will be

$$Z(t-1) + 2\tilde{u}\alpha/(\alpha - 1) = Z(t) + \tilde{u}\alpha/(\alpha - 1).$$

After $Y(t)$ has gone up n times in succession, in steps equal to $\tilde{u}$, the value of $Z(t+n-1) - Z(t-1)$ will approximately equal $n\tilde{u}\alpha/(\alpha - 1)$. Eventually, however, n will reach a value such that $Y(t+n-1) - Y(t-1) < n\tilde{u}$, which implies $X(t+n-1) - X(t-1) < n\tilde{u}$. The market price $Z(t+n-1)$ will then crash down to $X(t+n-1)$, losing all its excess growth in one swoop.

As the size of the original jump of X increases, the number of time intervals involved in smoothing also increases, and correction terms must be added.

Let us now discuss qualitatively the case where the value of the threshold $\tilde{u}$ is random. After a change of $Y(t)$, the market must attempt to determine whether it is a fully completed change of fundamental conditions, equal to a change of $X(t)$, or the beginning of a large change. In the first case, the motion need not "persist," but in the second case, it will. This naturally involves a test of statistical significance: A few changes of Y in the same direction may well "pass" as final moves, but a long run of rises should be interpreted as due to a "smoothed-off" large move. Thus, the following, more complicated pattern will replace the gradual rise followed by fast fall that was observed earlier. The first few changes of Z will equal the changes of Y, then Z will jump to such a value that its increase from the beginning of the rise equals $\alpha/(\alpha-1)$ times the increase of Y. Whenever the rise of Y stops, Z will fall to Y; whenever the rise of Y falters, and then resumes, Z will fall to Y and then jump up again.

In a further generalization, one may consider the case where large changes of Y are gradually transmitted with probability q and very rapidly transmitted with probability $1-q$. Then the distribution of the changes of Z will be a mixture of the distribution obtained in the previous argument and of the original distribution of changes of Y. The scaling character is preserved in such a mixture, as shown in M 1963e{E3}.

Remark on continuous-time processes. Let us return to the condition of discrete time made earlier in this subsection. Continuous-time processes with independent increments were considered in M 1963b{E14}. It was shown that in the L-stable case, $X(t)$ is significantly discontinuous, in the sense that if it changes greatly during a unit time increment, this change is mostly performed in one big step somewhere within that time. Therefore, the distribution of large jumps is practically identical to the distribution of large changes over finite increments. In the Gaussian case, to the contrary, the interpolated process is continuous. More generally, whenever the process $X(t)$ is interpolable to continuous time and its increments have a finite variance, there is a great difference between the distributions of its jumps (if any) and of its changes over finite time increments. This shows that the case of infinite variance – which in practice means the scaling case – is the only one for which the restriction to discrete time is not serious at all.

III.C. More complex economic models

There may be more than one "tracking" mechanism of the kind examined so far. It may, for example, happen that $Z(t)$ attempts to predict the future behavior of a smoothed-out form Y of $X(t)$, while $X(t)$ itself attempts to predict the future behavior of $\tilde{X}(t)$. This would lead to zigzags larger than those observed so far. Therefore, for the sake of stability, it will be very important in every case to examine the driving function $Y(t)$ with care: is it a smoothed-off fundamental economic quantity, or is it already influenced by forecasting speculation.

Suppose now that two functions $Z_1(t)$ and $Z_2(t)$ attempt to track each other (with lags in each case). The zigzags will become increasingly amplified, as in the divergent case of the cobweb phenomenon. All this hints at the difficulty of studying in detail the process by which "the market is made" through the interactions among a large number of traders. It also underlines the necessity of making a detailed study of the role that the SEL assigns to the specialist, which is to "insure the continuity of the market."

IV. COMMENTS ON THE VALUATION OF OIL FIELDS

Another illustration of the use of the results of Section IIB is provided by the example of oil fields in a previously unexplored country.

"Intuitively," there is a high probability that the total oil reserves in this country are very small; but, if it turns out to be oil-rich, its reserves would be very large. This means that the a priori distribution of the reserves X is likely to have a big "head" near $x=0$ and a long "tail"; indeed, the distribution is likely to be scaling. Let us now consider a forecaster who only knows the recognized reserves $Y(t)$ at time t. As long as the reserves are not completely explored, their expected market value $Z(t)$ should equal $\alpha Y(t)/(\alpha-1)$. The luckier the explorers have been in the past, the more they should be encouraged to continue digging and the more they should expect to pay for digging rights in the immediate neighborhood of a recognized source. Eventually, $Y(t)$ will reach X and will cease to increase; at this very point, $Z(t)$ will tumble down to $Y(t)$, thus decreasing by $Y(t)/(\alpha-1)$.

If the distribution of X were exponential, $Z(t)$ would exceed $Y(t)$ by an amount independent of $Y(t)$: the market value of the entirely unexplored territory. If $Y(t)$ were a truncated Gaussian variable, the premium for expected future findings would rapidly decrease with $1/Y(t)$.

It would be interesting to study actual forecasts in the light of those three possibilities. But the main reason for considering oil fields was to demonstrate how the variation of prices can be affected by *unavoidable* delays in the transmission of information about the physical world.

Acknowledgement

An earlier version, titled "Speculative Prices and a 'Martingale' Stochastic Model" was privately circulated in February, 1964. On December 29, 1964, it was presented at the Annual Winter Meeting of the Econometric Society, held in Chicago. I am greatly indebted to Professor Eugene F. Fama for many discussions that helped clarify the presentation. The present text also takes account of comments by Professor Charles Ying.

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

Editorial changes. The footnotes of the original were either incorporated in the text or placed as digressions at the end of subsections. A last section was moved to Chapter E3, for emphasis.

Annotation 6 Section IV. I discussed the valuation of oil fields in an IBM Report that was honored over thirty years later by coming out in print, as M 1995b. I understand that M1995b has become quite influential.