

Proof That Properly Anticipated Prices Fluctuate Randomly

Paul A. Samuelson
Massachusetts Institute
of Technology

The Enigma Posed

"In competitive markets there is a buyer for every seller. If one could be sure that a price will rise, it would have already risen." Arguments like this are used to deduce that competitive prices must display price changes over time, $X_{t+1} - X_t$, that perform a random walk with no predictable bias.

Is this a correct fact about well-organized wheat or other commodity markets? About stock exchange prices for equity shares? About futures markets for wheat or other commodities, as contrasted to the movement of actual "spot prices" for the concrete commodity.

Or is it merely an interesting (refutable) hypothesis about actual markets that can somehow be put to empirical testing?

Or is it a valid deduction (like the Pythagorean Theorem applicable to Euclidean triangles) whose truth is as immutable as $2 + 2 = 4$? Does its truth follow from the very definition of "free, competitive markets?" (If so, can there fail to exist in New York and London actual stock and commodity markets with those properties; and must any failures of the "truism" that turn up be attributable to "manipulation," "thinness of markets," or other market imperfections?)

The more one thinks about the problem, the more one wonders what it is that could be established by such abstract argumentation. Is the fact that American stocks have shown an average annual rise of more than 5 per cent over many decades compatible with the alleged "fair game" (or martingale property) of an unbiased random walk? Is it an exception that spot

wheat prices generally rise (presumably because of storage costs) from the July harvest time to the following spring and drop during June? Is the fact that the price of next July's future shows much less strong seasonal patterns a confirmation of the alleged truism? If so, what about the alleged Keynes-Hicks-Houthakker-Cootner pattern of "normal backwardation," in which next July's wheat future could be expected to rise in price a little from July harvest to, say, the following March (as a result of need of holders of the crop to coax out, at a cost, risk-disliking speculators with whom to make short-hedging transactions); and what about the Cootner pattern in which, once wheat stocks become low in March, processors wishing to be sure of having a minimum of wheat to process, seek short-selling speculators with whom to make long-hedging transactions, even at the cost of having the July quotation dropping a little in price in months like April and May?

Consideration of such prosaic and mundane facts raises doubt that there is anything much in celestial *a priori* reasoning from the axiom that what can be perceived about the future must already be "discounted" in current price quotations. Indeed, suppose that all the participants in actual markets are necessarily boobs when it comes to foreseeing the unforeseeable future. Why should "after-the-fact" price changes show *any* systematic pattern, such as non-bias? Are the very mathematical notions of probability of any relevance to actual market quotations? If so, how could we decide that this is indeed so?

Whatever the answers to these questions, I think we can suspect that there is no *a priori* necessity for actual Board of Trade grain prices to act in accordance with specific probability models. Perhaps it is a lucky accident, a boon from Mother Nature so to speak, that so many actual price time series do behave like uncorrelated or quasi-random walks. Thus, Maurice Kendall¹ almost proves too much when he finds negligible serial correlation in spot grain prices. For reasons that I shall discuss, we would not be too surprised to find this property in futures price changes. But surely spot prices ought to vary with shifts in such supply and demand factors as weather, crop yields, and crop plantings; or population, income, and taste changes. Who says that the weather must itself display no serial correlation? A dry month does tend to be followed by a dryer-than-average month because of persistence of pressure patterns, etc. Perhaps it is true that prices depend on a summation of so many small and somewhat independent sources of variation that the result is like a random walk. But there is no necessity for this. And the fact, if it is one, is not particularly related to perfect competition or market anticipations. For consider a monopolist who sells (or buys) at fixed price. If the demand (or supply) curve he faces is the resultant of numerous independent, additive sources of variation each of which is limited or small, his resulting quantity $\{q_t\}$ may well behave like a random walk, showing variations like the normal curve of error.

At this point, the reader may feel inclined to doubt that the arguments of my first paragraph have even a germ of interest for the economist. But I hope to show that such a rejection goes too far.

By positing a rather general stochastic model of price change, I shall deduce a fairly sweeping theorem in which next-period's price differences are shown to be uncorrelated with (if not completely independent of) previous period's price differences. This martingale property of zero expected capital gain will then be replaced by the slightly more general case of a constant mean percentage gain per unit time.

¹ See ref. [1].

You never get something for nothing. From a nonempirical base of axioms you never get empirical results. Deductive analysis cannot determine whether the empirical properties of the stochastic model I posit come at all close to resembling the empirical determinants of today's real-world markets. That question I shall not here investigate. I shall be content if I can, for once, find definite and unambiguous content to the arguments of the opening paragraph—arguments which have long haunted economists' discussions of competitive markets.

A General Stochastic Model of Price

Let $\{ \dots, X_{t-1}, X_t, X_{t+1}, \dots, X_{t+T}, \dots \}$ represent the time sequence of prices—as for example the price of spot #2 wheat in Chicago (or it could be a vector of prices or even quantities of several different goods). Given knowledge of today's price and of past prices $[X_t, X_{t-1}, \dots]$, suppose we cannot know with certainty tomorrow's price X_{t+1} , or any future price X_{t+T} . Suppose there is at best a probability distribution for any future price, whose form depends solely on the number of periods ahead over which we are trying to forecast prices, given by

$$(1) \quad \text{Prob}\{X_{t+T} = X | X_t = x_0, X_{t-1} = x_1, \dots\} \\ = P(X, x_0, x_1, \dots; T)$$

These P 's are assumed not to depend on, or change with, historical calendar time. In that sense, I posit a "stationary" process (but one consistent with, say, chronic inflation).

From the fundamental logic of probability, we can relate these probabilities relevant to different future forecast spans: $T = 1, 2, \dots$. All are ultimately deducible from the basic one-period $P(X, x_0, x_1, \dots; 1)$. Thus, always

$$\begin{aligned} P(X, x_0, x_1, \dots; 2) \\ &= \int_{-\infty}^{\infty} P(X, y, x_0, x_1, \dots; 1) \\ &\quad dP(y, x_0, x_1, \dots; 1) \\ (2) \quad &\dots\dots\dots \\ P(X, x_0, x_1, \dots; T) \\ &= \int_{-\infty}^{\infty} P(X, y, x_0, x_1, \dots; T-1) \\ &\quad dP(y, x_0, x_1, \dots; 1). \end{aligned}$$

These general Stieltjes integrals include summation of finite (or countably-infinite) probabilities and ordinary integrals of probability densities; no real ambiguity concerning the variable over which dP is being integrated will usually arise.

In words, (2) merely says that the probability of an event's happening is the sum of the probabilities of the different mutually-exclusive ways by which it could happen. If $P(X, x_0, x_1, \dots; 1)$ had the Markov property of being quite independent of (x_1, x_2, \dots) , (2) would be the so-called Chapman-Kolmogorov equation. But I do not assume any special Markov property. The generality of (1) and (2) must be emphasized. Nothing necessarily Gaussian or normal is assumed about any $P(X, x_0, x_1, \dots; T)$. It is possible, but not necessarily assumed, that an ergodic state for P will emerge in the limit as T goes to infinity. Thus,

$$(3) \quad \lim_{T \rightarrow \infty} P(X, x_0, x_1, \dots; T) = P(X) \text{ independently of } (x_0, x_1, \dots)$$

would involve such an ergodic property.

Here are some examples of possible processes that define the P 's. Suppose $\{X_t\}$ satisfies a Yule-Wold autoregressive linear equation of the type

$$(4) \quad X_{t+1} = aX_t + \{u_t\}, \quad |a| < 1,$$

with $\{u_t\}$ independent drawings from the same table of random digits or of Gaussian variates.

Or suppose the X_t price can take on only the finite discrete values $[Q_1, Q_2, \dots, Q_n]$; and let a_{ij} , the nonnegative coefficients of a Markov transitional-probability matrix, represent the probability that a price now observed to be Q_i will one period later be observed to be Q_j , where $\sum_j a_{ij} = 1$ for all i . If $[a_{ij}] = A$, the reader can verify that A^2 defines $P(X, x_0; 2)$ and A^T defines $P(X, x_0; T)$. If all $a_{ij} > 0$, A is a well-defined ergodic state, composed of the matrix $[a_{ij}]$ with identical rows and hence independent of the initial observed Q_i value.

As a third possible model let $\{X_t\}$ take a random walk in the sense of doubling in the next period with probability $1/3$ and halving with probability $2/3$. This defines $P(X, x_0; 1)$, and the reader can verify that (2) gives for $P(X, x_0; 2)$ the property that

X_{t+2}/X_t will be $(4, 1, 1/4)$ with respective probabilities $(1/9, 4/9, 4/9)$, and so on with the usual binomial distribution for each T and with the central-limit theorem showing that a normal distribution is approached for $\log(X_{t+T}/X_t)$ as $T \rightarrow \infty$. This is an instance of a multiplicative Brownian motion applied to prices. Unlike the absolute or additive Brownian motion, it has the grace to avoid negative prices.

Specification of a Model Defining Behavior of a Futures Price

Now consider today's "futures price quotation" for the actual spot price that will prevail T periods from now—i.e., the price quoted today at t for a contract requiring delivery of actual physical goods at time $t + T$. If the present time is t , with present spot price X_t , the relevant spot price that is to prevail later is given by X_{t+T} . The newly defined futures price, quoted today, for that future X_{t+T} , we might denote by $Y(T, t)$. When another period passes, we shall know $(X_{t+1}, X_t, X_{t-1}, \dots)$ instead of merely (X_t, X_{t-1}, \dots) ; and the new quotation for the same futures price we have been talking about will be written as $Y(T-1, t+1)$. It in turn will be succeeded by the sequence $[Y(T-2, t+2), \dots, Y(t-n, t+n), \dots, Y(1, t+T-1), Y(0, t+T)]$. After $t+T$, there is no problem of pricing this particular futures contract. Thus, the July, 1964, Chicago Wheat Contract for delivery of wheat became closed, ancient history after July, 1964; but for more than twelve months prior to that date, its quotation oscillated from period to period and could be read in the newspaper.

What relationship shall we posit between the sequence $\{Y(T-n, t+n)\}$ and the sequence $\{X_{t+1}\}$? When the due date for the futures contract arrives, arbitrage will ensure that

$$(5) \quad Y(0, t+T) = X_{t+T}, \text{ commissions aside.}$$

A period earlier no one can know what X_{t+T} will turn out to be. If interest and risk-aversion can be ignored, it is tempting to assume that people in the market place make as full use as they can of the posited probability distribution $P(X_{t+T}, X_{t+T-1}, X_{t+T-2}, \dots; 1)$ of next-period's price and bid by supply and demand $Y(1, t+T-1)$

to the mean or mathematically-expected level of tomorrow's price. That way neither short-sellers nor long-buyers stand to make a positive gain or loss. This constitutes the rationale of my first model of futures price, which is based on the following.

Axiom of Mathematically Expected Price Formation. If spot prices $\{X_t\}$ are subject to the probability distribution of (1), a futures price is to be set by competitive bidding at the now-expected level of the terminal spot price. That is,

$$(6) \quad Y(T, t + T) = E[X_{t+T} | X_t, X_{t-1}, \dots], \\ (T = 1, 2, \dots) \\ = \int_{-\infty}^{\infty} X dP(X, X_t, X_{t-1}, \dots; T).$$

The Basic Theorem

Equations (1) and (6) completely determine the properties of the model. I can now derive from them the basic theorem at which I hinted in the usual vague arguments expressed in the opening paragraph. Let us observe numerous sequences of futures prices generated by this model, up until their terminal data. They will turn out, *on the average*, to have no upward or downward drift anywhere! (This will be true, regardless of the systematic seasonal patterns in X_t . This would be true, under our axiom, even in time of severe inflation or deflation of X_t itself—only remember, please, that the neglect of interest underlying our axiom would seem unwarranted in time of confidently-anticipated extreme inflation.)

Theorem of Fair-Game Futures Pricing. If spot prices $\{X_t\}$ are subject to the probability laws of (1) and (2), and the futures price sequence $\{Y(T, t + T), Y(T - 1, t + T - 1), \dots, Y(1, t + T - 1), Y(0, t + T)\}$ is subject to the axiom of expected price as formulated in (5) and (6), then the latter sequence is a fair game (or martingale) in the sense of having unbiased price changes, or

$$(7) \quad E[Y(T - 1, t + T - 1) | X_t, X_{t-1}, \dots] \\ \equiv Y(T, t), \quad \text{or}$$

writing

$$(8) \quad Y(T - 1, t + T - 1) - Y(T, t) = \Delta Y(T, t), \\ E[\Delta Y(T, t)] \equiv 0;$$

from which it follows inductively that

$$(9) \quad E[\Delta^n Y(T, t)] \equiv 0 \quad (n = 1, 2, \dots, T).$$

This means that there is no way of making an expected profit by extrapolating past changes in the futures price, by chart or any other esoteric devices of magic or mathematics. The market quotation $Y(T, t)$ already contains in itself all that can be known about the future and in that sense has discounted future contingencies as much as is humanly possible (or inhumanly possible within the axiom of the model).

The theorem does not imply that the sequence of Y 's perform a Brownian motion. It does not imply that $\Delta Y(T, t)$ is statistically *independent* of $\Delta Y(T + 1, t - 1)$; it implies only that given knowledge of $Y(T, t)$ the Pearsonian correlation coefficient between the above two Δ 's will be zero. It is a source of comfort to the economist, rather than otherwise, that wheat prices should not perform a Brownian random walk. A Brownian walk, like the walk of a drunken sailor, wanders indefinitely far, listing with the wind.

Surely, economic law tells us that the price of wheat—whether it be spot $\{X_t\}$ or futures $\{Y(T, t)\}$ —cannot drift sky-high or ground-low. It does have a rendezvous with its destiny of supply and demand, albeit our knowledge of future supply and demand trends becomes dimmer as the envisaged date recedes farther into the future. We would expect people in the market place, in pursuit of avid and intelligent self-interest, to take account of those elements of future events that in a probability sense may be discerned to be casting their shadows before them. (Because past events cast *their* shadows after them, future events can be said to cast their shadows before them.)

Although the sequence $\{\Delta Y(T - n, t + n)\}$ has a zero first moment at all time periods $T - n$, there is no reason to suppose that the riskiness of holding a futures—in the sense of the second moment of variance, as measured by $E[\{\Delta Y(T - n, t + n)\}]$ —should be the same when T is large and the terminal date far away as when $T - n$ is small and the futures contract about to expire. It is a well-known rule of thumb that nearness to expiration date involves greater variability or riskiness per hour or per day or per month than does farness. Partly this is due, I think, to factors not encompassed in the present model—for example, the real-life complications

that make arbitrage equality shown in (5) hold only approximately. However, the present theory can contribute an elegant explanation of why we should expect far-distant futures to move more sluggishly than near ones. Its explanation does not lean at all on the undoubted fact that, during certain pre-harvest periods when stocks are normally low, changes in spot prices $\{\Delta X_t\}$ can themselves be expected to experience great volatility and second moment. Instead, it uses the contrary minor premise of posited *uniformity* through time of the distribution of $\{\Delta X_t\}$ —so that it constitutes a stationary time series—to deduce the law of increasing volatility of a maturing futures contract. Before deriving the result, I present a proof of the martingale theorem.

By definition

$$Y(T, t) = \int_{-\infty}^{\infty} X dP(X, X_t, X_{t-1}, \dots; T) ,$$

axiom of (6)

$$Y(T-1, t+1) = \int_{-\infty}^{\infty} X dP(X, X_{t+1}, X_t, X_{t-1}, \dots; T-1) ,$$

axiom of (6)

$$\begin{aligned} &= f(X_{t+1}, X_t, X_{t-1}, \dots) \\ E[Y(T-1, t+1) | X_t, X_{t-1}, \dots] &= \int_{-\infty}^{\infty} f(Z, X_t, X_{t-1}, \dots) dP(Z, X_t, X_{t-1}, \dots; 1) \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} X dP(X, Z, X_t, \dots; T-1) \right] \\ &\quad dP(Z, X_t, \dots; 1) \\ (10) &= \int_{-\infty}^{\infty} X d \left[\int_{-\infty}^{\infty} P(X, Z, X_t, \dots; T-1) \right. \\ &\quad \left. dP(Z, X_t, \dots; 1) \right] \end{aligned}$$

$$\begin{aligned} &= \int_{-\infty}^{\infty} X dP(X, X_t, \dots; T) , \text{ by (2)} \\ &= Y(T, t) \quad \text{Q.E.D.} \end{aligned}$$

In thus proving (7), I have made permissible interchanges in the order of integration of the double Stieltjes integrals.

The content of this general theorem can be illustrated by some specific stochastic processes.

(a) Thus, suppose $\{X_t\}$ represents a simple Brownian motion without bias. Then $Y(T, t)$ becomes nothing but X_t it-

self. Since $\{X_t\}$ is a fair-game with $E[\Delta X] = 0$, then so must $E[\Delta Y] = 0$.

(b) Going beyond this simple case, suppose $\{X_t\}$ is a random walk with biased drift so that $E[\Delta X] = \mu \neq 0$. Then $Y(T, t) = X_t + \mu T$ and $E[\Delta Y] = E[\Delta X] + \mu(T-1-T) = \mu - \mu = 0$, as the theorem requires.

(c) Now let $\{X_t\}$ satisfy an autoregressive stochastic relation

$$X_{t+1} = aX_t + \{u_t\} , \quad E(u_t) = \mu ,$$

where $\{u_t\}$ is an unserially correlated random variable with distribution $P(U)$. Then

$$\begin{aligned} Y(T, t) &= a^T X_t + \mu T \\ E[\Delta Y] &= \int_{-\infty}^{\infty} a^{T-1} (aX_t + U_t) dP(U_t) \\ &\quad + \mu(T-1) - a^T X_t - \mu T \\ &= 0 + \mu - \mu = 0. \end{aligned}$$

(d) Finally, let $A = [a(i, j)]$ be a finite Markov transition-probability matrix giving

$$\text{Prob}\{X_{t+1} = Q_i | X_t = Q_j\}.$$

Write $A^T = [a_{ij}^T]$, and

$$\begin{aligned} Y(T, t) &= [Q_1, \dots, Q_n] \text{ times the } j^{\text{th}} \text{ column of } A^T, \text{ by Axiom} \\ &= \sum_i Q_i a_{ij}^T, \text{ denoted by} \\ &\quad Y(T | Q_j). \end{aligned}$$

Then

$$\begin{aligned} E[Y(T-1), t+1 | X_t = Q_j] &= \sum_k Y(T-1 | Q_k) a_{kj}, \\ &\quad \text{by definition} \\ &= \sum_k Q_k a_{kj}^{T-1} a_{kj} \\ &= \sum_k Q_k a_{kj}^T, \text{ since } A^{T-1} A = A^T \\ &= Y(T, t) \quad \text{Q.E.D.} \end{aligned}$$

The theorem is so general that I must confess to having oscillated over the years in my own mind between regarding it as trivially obvious (and almost trivially vacuous) and regarding it as remarkably sweeping. Such perhaps is characteristic of basic results. And actually the empirical question of the applicability of the model to economic reality must be kept distinct from the logical problem of what is the model's implied content.

Figure 1 should help to explain the theorem. It supposes $\{X_t\}$ is generated by

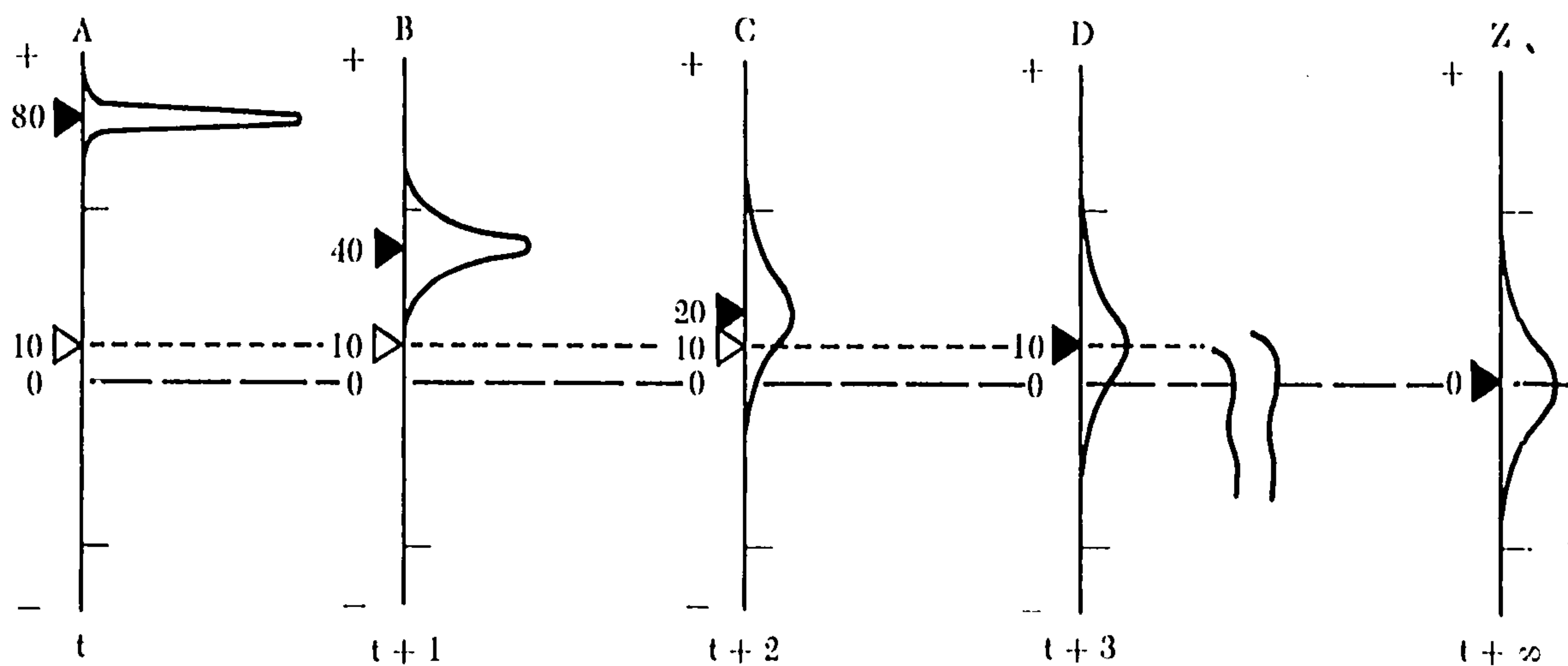


Figure 1A.

At time t , we know $X_t = 80$ with certainty. We can expect, from (4) with $a = 1/2$ and $\{u_t\}$ approximately normally distributed, that X_{t+1} be distributed as shown, with $E[X_{t+1}|X_t = 80] = 40$; for X_{t+2} , X_{t+3} the distributions are becoming slightly more dispersed, with means drifting downward like $80(2^{-2})$, $80(2^{-3})$; because $a = 1/2 < 1$, there is a limiting ergodic state for $t + \infty$, much as shown, with $E[X_{t+\infty}|X_t] = 0$ independently of X_t . At time t , the futures price X_{t+3} is evaluated at 10, as shown by the white arrow at A, which merely reflects what can now be known about the distribution and mean of the spot price three periods from now. Now ask

yourself the question, What do I think at this time t will be the likely distribution of this futures price $Y(3, t)$ at the next period when it will have become $Y(2, t + 1)$? It will be distributed around $1/4$ of next period's spot price X_{t+1} in exactly the same way that C shows X_{t+2} distributed around $1/4$ of the period's spot price X_t . You do not today know where tomorrow's X_{t+1} will fall; but B does give you its probability distribution. Combining these bits of information, Figure 1B below plots (with enlarged vertical scale) what you *now* are entitled to regard as the probability distribution for $Y(3, t)$, $Y(2, t + 1)$, $Y(1, t + 2)$, $Y(0, t + 3) = X_{t+3}$.

(4)'s autoregressive model with $a = 1/2$. Random shocks $\{u_t\}$ aside, if price is once perturbed above (or below) its normal value (set at zero by convention), in each period it will return one-half way back to zero.

Generalizing the Theorem

If money has to be tied up in holding the $Y(T, t)$ contract and if there is a positive safe rate of interest, the process—being a fair game—does not stand to earn even the opportunity-cost of foregone safe interest. If, in addition, people have risk aversion (so that the utility whose expected value they seek to maximize is a strictly concave function of money wealth or income), the contract will probably have to promise a positive percentage yield per unit time, R , where $1 + R = \lambda > 1 + \text{safe interest}$.

Therefore, we replace the simple axiom of mean expected value by a slightly more general axiom.

Axiom of Present-Discounted Expected Value: At each point of time t for a future price, we posit

$$Y(T, t) = \lambda^{-T} E[X_{t,T}|X_t, X_{t-1}, \dots];$$

or, slightly more generally, since the variability of Y may be slightly different for each value of T , we may expect a different $(\lambda_1, \lambda_2, \dots, \lambda_T)$ yield for which people hold out if they are to hold for the next period a futures contract with T years to go. The present-discounted expectation becomes

$$Y(T, t) = \lambda_1^{-1} \lambda_2^{-1} \dots \lambda_T^{-1} E[X_{t,T}|X_t, X_{t-1}, \dots].$$

Just as the simple axiom of expected gain led to the theorem of unbiased price change, the new general axiom leads to the theorem on “normal backwardation”.

Theorem of Mean Percentage Price Drift. If spot prices $\{X_t\}$ are determined by the stipulated general stochastic process $P(X_{t,T}, X_t, X_{t-1}, \dots; T)$ and we define the sequence

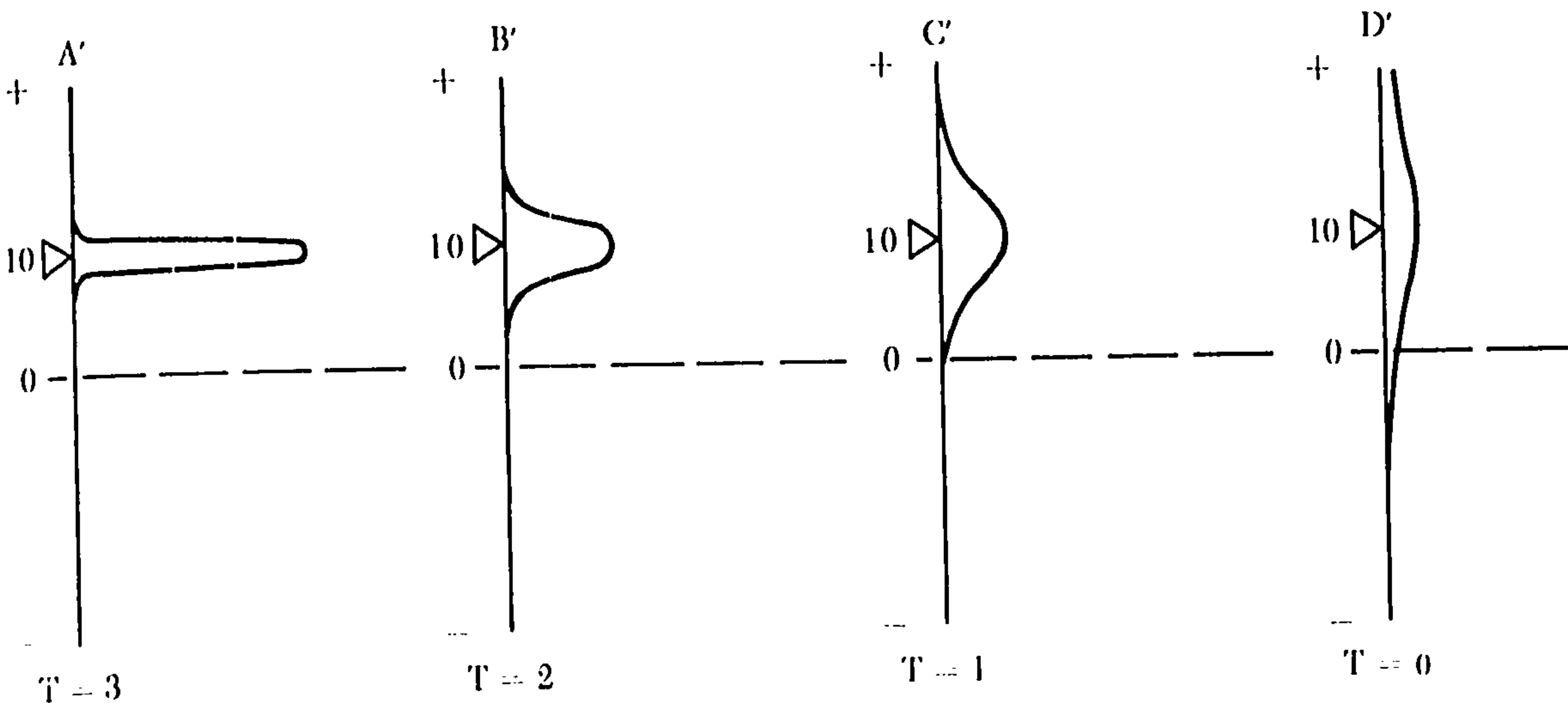


Figure 1B.
 A', B', C', and D' give the probability distributions you are entitled to envisage on the basis of today's certain knowledge of $[X_t, X_{t+1}, \dots$ and $Y(3, t)]$ for next periods' $[Y(2, t + 1), Y(1, t + 2), Y(0, t + 3) = X_{t+3}]$. Note that all have the same mean of 10 but that your cone of uncertainty widens as the time in which new unknown, independent disturbances $\{u_{t,i}\}$ can intervene. Note that after one period passes, both diagrams will have to be redrawn: A and A' will be irrelevant; a particular point on B and B' will become the new present, and the new (C, C') and (D, D') will bear the same relation to (B, B') that now (C, C') and (B, B') bear respectively to (A, A'). If the stochastic variables $u_{t,1}, u_{t,2}$, and $u_{t,3}$ were identically zero, the futures price would show no variability, staying always at 10. $Y(3,t)$ and $Y(2, t + 1)$ are likely to be more alike with a less volatile difference ΔY than can $Y(2, t + 1)$ and $Y(1, t + 2)$. (For equation (4), the variance of $\Delta Y(t - n, t + n)$ is proportional to a^{T-n} , diminishing as $T - n$ becomes large.) Why? Because the first pair stand to have $(u_{t,2}, u_{t,3})$ in common and differ only in that the first of the pair also has u_t as a source of variation.

The fact that this pair has two out of three elements of variation in common is stabilizing on their difference ΔY in comparison with the second pair which have only $u_{t,3}$ out of the possible $(u_{t,2}, u_{t,3})$ in common. A far-distant future will not change much in the next month since so few of the disturbances upon which its fate depends will change in this month; it stays close to the general level given by the so-called law of averages. At the other extreme, $Y(1, t + 2)$ and $Y(0, t + 3) = X_{t+3}$ will differ because of the single unknowable $u_{t,3}$, and this sudden-death kind of situation will subject their difference to great variance without any cushioning from the law of large numbers. A striking way of seeing this is to disregard the varying means and to think of the passage of one period as leaving us at the left of the diagram subtracting the last (rather than the first, as when we shift the present from A to B) of the vertical axes. Precisely because D approaches an ergodic state, C and D differ less than do A and B in Figure 1A. So losing the last axis, which is like making C become the new D, gives less of a change than would shifting B to A.

$$\{Y(T,t), Y(T-1, t+1), \dots, Y(1, t+T-1), Y(0, t+T)\} \text{ by}$$

$$Y(T, t) = \lambda_T^{-1} \dots \lambda_1^{-1} E[X_{t+T} | X_t, X_{t+1}, \dots]$$

$$= \lambda_T^{-1} \dots \lambda_1^{-1} \int_{-\infty}^{\infty} X dP(X, X_t, X_{t+1}, \dots; T),$$

it follows that

$$E[Y(T-n, t+n) | X_t, X_{t+1}, \dots] = \lambda_T \dots \lambda_{T-n+1} Y(T, t) = \lambda^n Y(T, t) \text{ if } \lambda_i \equiv \lambda.$$

In words, this says that the futures price will rise in each period by the percentage

$\lambda - 1$ (or $\lambda_T - 1$, where presumably $\lambda_1 > \lambda_2 > \dots > \lambda_T$ if people abhor the extra riskiness inherent in a futures contract as it matures). This theorem provides rational account of the Keynes-Houthakker-Cootner² doctrine of "normal backwardation". It says that, within the defined model, all chart methods attempting to read out of the past sequence of known prices $[X_t, X_{t+1}, \dots, Y(T,t), Y(T+1, t-1), \dots]$ any profitable pattern of prediction is doomed to failure. So to speak, the market has already,

² See refs. [2], [3], [4], [5], and [6].

by our axiom, discounted all knowable future information so that the present-discounted variable $(\lambda_T \dots \lambda_1)^{-1} Y$ sequence is itself a fair-game martingale.

Re-examining the proof of the simple theorem, we see that the same interchange of order of integration in the relevant double integral gives an immediate proof of the theorem.

Finally, examining the proof shows that a still more general theorem is valid. Wherever we have had $X_{t,T}$ in $E[X_{t,T}|X_t, X_{t-1}, \dots]$ or X in the double integral, we could have put in any function of X , say $g(X)$ or $g(X_{t,T})$ and still preserved the martingale properties. This is important because, as mentioned, we may want to interpret X_t as a vector—as e.g. with components (soft wheat price, hard wheat price, corn price, midwest rainfall, national income, etc.). Then the scalar $g(X_{t,T})$ might be the selection of soft wheat price, the deliverable goods for the futures contract in question. Only a deluded chartist would think that extraneous variables, such as rainfall, can be excluded from an optimal *probability* description of future soft wheat prices; and yet it is true that, after speculators have taken such variables into account in implementing the axiom of expected gain, the resulting Y sequence tells its own (simple!) story.

Still one more bargain in cheap generality can be garnered. As far as the theorem's proof is concerned, $g(X_{t,T})$ could just as well be a function also of all present known data and be written as $g(X_{t,T}, X_t, X_{t-1}, \dots)$, without affecting its general martingale property. It would even be possible to put now-known values of $Y(T+k, t-k)$ as variables in g , but there would seem to be no need to do this since ultimately all known Y 's should be reducible to functions of the X_{t-k} inputs. Still there is no harm in so including the Y 's, and in some cases the form of the function might be much simplified by including such Y 's (and even the Y 's of other terminal date). To illustrate how one might want X_t in $g(X_{t,T}, X_t, \dots)$, suppose that either number 1 or number 2 grades of soft wheat are deliverable on the July, 1965, contract. Let the vector X_t contain as elements price data on both grades; and suppose that if one of these is now known to be much cheaper than the other, one can be pretty sure that this cheaper

grade will be the grade actually delivered on the contract at the terminal date; and hence for certain value of the elements of the vector X_t , $g(\dots)$ will be the expected price of one grade rather than of the other.

Conclusion

A result of some generality has now been established. Anyone who thinks it obvious should reflect on the following fact: if instead of taking an expected value or mean value for $X_{t,T}$, we set $Y(T,t)$ at, say, its median value, then it will not necessarily be true that expected value of ΔY is zero. Nor is it true that the median of ΔY so defined is zero. For example, let $P(X_{t+1}, X_t, 1) = P(X_{t+1}/X_t)$ with $P(1) > 1/2 = P(m)$, $0 < m < 1$. Then $Y(T,t) = mX_t$, $Y(T+1, t+1) = mX_{t+1}$. For median $\Delta Y = 0$, we require

$$\begin{aligned} \text{Prob}\{Y_{t+1} \leq Y_t\} &= \text{Prob}\{mX_{t+1} \leq mX_t\} \\ &= \text{Prob}\{X_{t+1} - X_t\} = \frac{1}{2}. \end{aligned}$$

But actually,

$$\begin{aligned} \text{Prob}\{X_{t+1} \leq X_t\} &= \\ \text{Prob}\left\{\frac{X_{t+1}}{X_t} \leq 1\right\} &= \\ &P(1) > \frac{1}{2}. \end{aligned}$$

Hence, ΔY is more likely to be negative than positive—even if $P(X/x)$ defines a fair game with $E(X/x) = x$.

One should not read too much into the established theorem. It does not prove that actual competitive markets work well. It does not say that speculation is a good thing or that randomness of price changes would be a good thing. It does not prove that anyone who makes money in speculation is *ipso facto* deserving of the gain or even that he has accomplished something good for society or for anyone but himself. All or none of these may be true, but that would require a different investigation.

I have not here discussed where the basic probability distributions are supposed to come from. In whose minds are they *ex ante*? Is there any *ex post* validation of them? Are they supposed to belong to the market as a whole? And what does

that mean? Are they supposed to belong to the "representative individual," and who is he? Are they some defensible or necessitous compromise of divergent expectation patterns? Do price quotations somehow produce a Pareto-optimal configuration of *ex ante* subjective probabilities? This paper has not attempted to pronounce on these interesting questions.

References

1 Kendall, Maurice G., "The Analysis of Economic Time-Series—Part I: Prices," *Journal of the Royal Statistical Society* 96, Part I (1953), pp. 11-25; also pp. 85-99 in Cootner, Paul H.

(editor), *The Random Character of Stock Market Prices*. Cambridge: The M.I.T. Press, 1964.

2 Keynes, J. M., *A Treatise on Money*, Vol. II: *The Applied Theory of Money*. London: Macmillan & Company, 1930.

3 Houthakker, H., "The Scope and Limits of Futures Trading," pp. 134-159 in Abramovitz, Moses (editor), *Allocation of Economic Resources*.

4 Houthakker, H., "Systematic and Random Elements in Short-Term Price Movements," *American Economic Review* 51 (1961), pp. 164-172.

5 Cootner, Paul H., "Returns to Speculators: Telser vs. Keynes," *Journal of Political Economy* 68 (1960), pp. 396-404.

6 Cootner, Paul H., "Rejoinder," *Journal of Political Economy* 68 (1960), pp. 415-418.