PERIODIC STRUCTURE IN THE BROWNIAN MOTION OF STOCK PRICES

M. F. M. Osborne

U. S. Naval Research Laboratory, Washington, D. C.

(Received January 2, 1962)

The internal structure of stock prices is examined by comparison with simple random walks of basic step $\frac{1}{8}$, in which the individual price changes ΔP are the step length, and the volume measures the rate at which the steps are taken. It is found that there is definite evidence of periodic in time structure corresponding to intervals of a day, week, quarter, and year; these being simply the cycles of human attention span. The evidence is not in the periodicity of the price sequences $P(t)$, rather in the distribution of the first and second differences of $P(t)$, especially the second moment of ΔP (or variance), and in the rate at which the steps are taken. It is also shown that there is a periodic 'space structure' in the price coordinate P, corresponding to the Brownian motion in the presence of equally spaced sites of preferred occupancy and reflection barriers, at the whole numbers. There is also marked evidence of 'clustered' activity, the data being analyzed by methods appropriate to cosmic ray bursts, or star counts on astronomical photographs. In general, the picture of price motion as simple random walks is supported qualitatively; quantitatively there are some substantial departures from this simple picture.

IT IS THE purpose of this paper to examine some internal properties of common stock prices in detail, in order to see more precisely how the general properties of Brownian motion previously reported[1,2] arise. The basic material of our discussion is the ensemble of sequences of numbers (prices) in time, $P_j(t)$, where the index j runs over the members of the ensemble, or individual common stocks, and t is the time at which the transaction giving the price, $P_j(t)$, is executed. We shall also consider the ensemble of sequences of volume in round lots, $V_j(t,\tau)$. Here, we must also specify the interval, τ, of time to which the volume refers, whether it is a single block of transactions at one instant, ($\tau=0$) on the tape, or the sum of all transactions for intervals τ of a day, week, or month.

Now neither the set of numbers $P_j(t)$ nor $V_j(t,\tau)$ is in fact normally distributed, whether one forms a set by holding j fixed to a single stock and letting t vary, or conversely. But the following derived sets are normally distributed, approximately.[1]

1. $\log_e P_j(t)$ $\quad$ t fixed, j varying.

2. $\Delta(\tau)\log_e P_j(t)=\log_e[P_j(t+\tau)/P_j(t)]$ for either (a) τ, t fixed, j varying (b) j, τ fixed, $t=t_0$, $t_0+\tau$, $t_0+2\tau$, $t_0+3\tau$, etc.

We hope to show also that $\log_e V_j(t,\tau)$ for either t, τ fixed, j varying as in 2*a* or j, τ fixed, t varying at intervals τ as in 2*b*, is approximately normally distributed, provided τ, the interval over which the volume is summed, is one day. For longer values of τ (week, month), one might expect from the central limit theorem that the volume would at least for the second case (like 2*b*), approach a normal rather than a log-normal distribution. We shall discuss whether this does in fact occur.

From either the price or volume ensemble of sequences, we can define two types of distributions. An 'across-the-market' frequency distribution is for a set in which only the stock index j is allowed to vary (cases 1 or 2*a* above). A 'sequential distribution' is for a set for which the members form a sequence in time (2b) and for which the index j is held fixed. Dispersions from these two distributions are referred to respectively as across the market or sequential. In the notation of reference 1, $\sigma_S = \sigma[\Delta(\tau)\log_e P]_t$ was an across-the-market dispersion; whereas, σ_M, for market index changes, and $\sigma_Y = \sqrt{\sigma_S{}^2 + \sigma_M{}^2} = \sigma[\Delta(\tau)\log_e P]_j$ were sequential dispersions. An across-the-market dispersion for an interval τ can be determined from data from one interval τ. A sequential dispersion for an interval τ requires N nonoverlapping intervals of length τ, where N is the number of separate observations in the data sample.

To estimate either a sequential or across-the-market dispersion we can use one of two methods. If the number in the sample is 'large' (and in practice 12 or more will be considered 'large'), we will use one-half the intersextile range as our estimate; i.e., one-half the range from the upper to lower one-sixth of our sample. This is, in fact, a good estimate of dispersion if the sample is withdrawn from a normal population. If not, one simply understands this estimate as one-half the range around the median containing two-thirds of the data points.

For samples 'small' in number (3, 4, 5) we will simply take the range from the largest to the smallest, times a correction factor determined under the assumption that the parent population is in fact normally distributed.[3] For small samples from a normal population, the range is a fairly efficient estimator of dispersion.

We can also summarize here the following statistical facts that we shall frequently see. If we have a variable y that is log-normally distributed, then $Z = \log y$ is normally distributed and the log of geometric mean of y is the arithmetic mean Z. The log of median y is the median Z, close to the arithmetic mean Z, if Z is normal, or y log-normal. If Z is a normal variable, of median zero and variance σ^2, then for the log-normal variable y, $\bar{y} - y_{\text{median}} = e^{\sigma^2/2} - 1$, a difference considered negligible unless $\sigma \gtrsim 1$.

A second point to note is that if we have a variable y whose probability density distribution is $\varphi(y)\, dy$ (normal or not), then the variance of the variable $Z=f(y)$ is approximately $\sigma_Z^2/Z=\sigma_y^2|\partial f/\partial y|^2_{y=\bar{y}}$, if one term of a Taylor series expansion for $f(y)$ is a fair approximation over a range $y=\bar{y}\pm\sigma_y$. If we apply this approximate rule to the particular case of $Z=\log y$, we find $\sigma(Z)=\sigma(\log y)\cong\sigma_y/\bar{y}$.

We shall have frequent use for χ^2-tests as measures of significance[3,13] and shall also have occasion to employ the following rule of thumb based on the binomial distribution: The dispersion of the expected number in a class interval of a histogram is approximately the square root of that expected number.

STOCK PRICES AS SIMPLE RANDOM WALKS

WITH THIS background we can now examine some data. In reference 1 we showed that the observed across-the-market dispersion developed in $\Delta(\tau)\log_e P$ for $\tau=$one day, $\sigma[\Delta(\text{day})\log_e P]_t=\sigma_S(\text{day})$, was in order of magnitude agreement with that computed for a typically-priced (\$40) stock, whose 'average' volume = total daily market volume/number of issues traded, was 10 to 20 round lots per day, with 1 to 2 round lots per transaction, and $\pm\$1/8$ the step length. Figure 1 gives the actual number of transactions plotted against the volume for a number of different stocks for a given day, and it will be seen that the volume V is indeed a good measure of the number of transactions T. The line $V=1.5T$ fits the data quite well, as we originally surmised.

While our original picture of a median priced stock and an 'average' volume seems to fit our data quantitatively, a little reflection will indicate that it cannot possibly be correct in detail. The dispersion in position at the end of T steps for a simple coin-tossing random walk, of step length h is $\sigma=\sqrt{T}h=\sqrt{(V/1.5)}h$. If we apply this expression to stock prices, where $h=1/8$ point is the step length, then we have the following two conclusions, which we can test by comparison with data.

(1) Suppose we compare the across-the-market dispersion of two groups of stocks, each member of which trades in approximately the same volume. One group starts in a price range \$20–\$30, the other starts in the range \$80–\$120. Then if the above formula for σ is correct, the across-the-market dispersion of changes in $\log_e$ price for the low-priced group will be approximately four times that for the high-priced group. For this restricted experiment with ΔP denoting the change of price after T steps, we have

$$\sigma(\Delta\log_e P)\simeq\sigma(\Delta P/P)\cong\sigma(\Delta P)/\bar{P}\cong h\sqrt{V/1.5}/P_{\text{initial}}. \qquad (1)$$

Figure 2, using data from the stocks listed in Table I, makes this comparison, and it will be seen that the low-priced stocks do indeed have a larger dispersion as predicted, but it is only 1.5 times as large as that for high-priced stocks—not four times as large as the simple theory would indicate. This observation is in agreement with the tradition of the market place that, other things being equal, low-priced stocks are percentagewise somewhat more volatile than high-priced; the quantitative agreement with the simple theory is not very satisfactory.

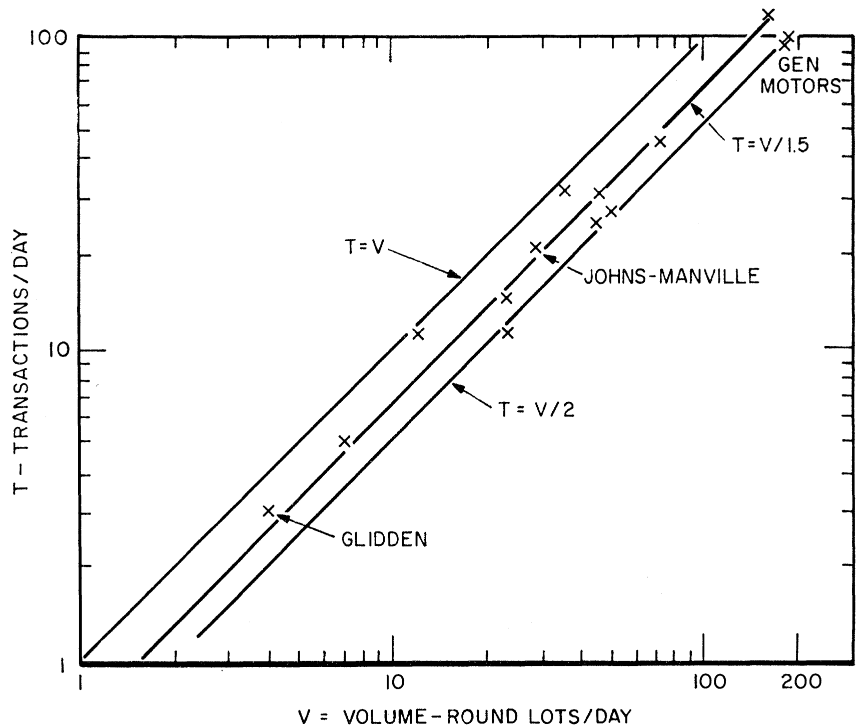

Fig. 1. Transactions per day vs. round lot volume per day, for October 30, 1959. Data from Fitch's "A Daily Market Quotation Publication," 138 Pearl St., New York 5, N. Y.

Price and 'quality' are sometimes considered to be correlated, and there is some evidence[4] that it is 'quality' rather than price which determines the dispersion. It is for this reason we have identified the stocks of Fig. 2 in Table I, so that the reader may judge for himself the relative quality of these samples of high- and low-priced stocks.

We have also plotted in Fig. 2 the theoretical curves for the average volume indicated at bottom of Table I, using 1.5 round lots per transaction. It will be seen that the theoretical dispersions are considerably larger than the observed for the low-priced stocks; the data on high-priced stocks are quite well represented. It will also be noted that the average volumes of

Table I are considerably larger than are average volumes computed by dividing the total market volume by the number of issues traded. For the first six months of 1960 this would have been about 27 round lots per day per issue.

Aside from the experimental evidence, one can put the objection to the simple picture of steps of $\frac{1}{8}$ point on purely theoretical grounds. Such

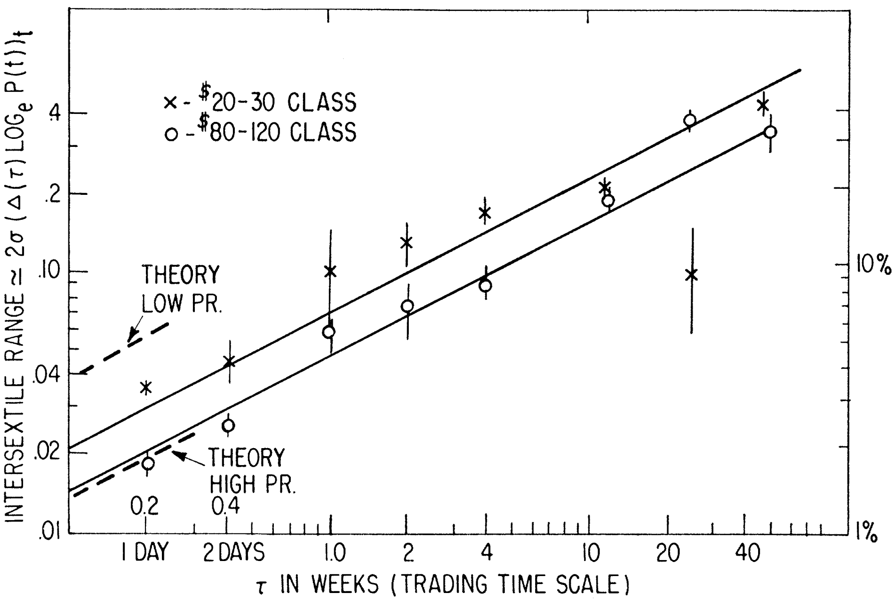

Fig. 2. Across-the-market dispersion (doubled) of $\Delta\log_e P$ as a function of time interval τ, for two samples of 12 NYSE common stocks each, in the \$20–\$30 and \$80–\$120 initial price range. Initial date January 4, 1960. Data taken graphically from charts of Trader's Research, Inc., Lambert Airport, Missouri. The denoted range of uncertainty is whichever was the larger interval, from 2nd to 3rd or 10th to 11th member of the sample. The reason for the anomalous behavior of the data at 24 weeks is unknown.

behavior would contradict a previous assumption in reference 1. If this simple picture were completely correct, the percentage risk (as measured by dispersion of $\Delta\log_e P$ for constant volume) would vary inversely as the price, whereas in fact the percentage risks, or dispersion of $\Delta\log_e P$ in high- or low-priced stocks are *nearly* constant (to a factor 1.5), more nearly in agreement with what we assumed for risk-taking in the face of values determined by the Weber-Fechner law.

(2) A second prediction that the simple theory makes is that the dispersion of price changes should increase with the volume, more particularly as $\sqrt{V/1.5}$. Which dispersion (sequential or across the market) and what is the proper measure of volume is subject to interpretation.

One interpretation and test of the above statement using readily available data, is to plot the across-the-market dispersion of 1000 or more *common* stocks, taken from the histograms of "The Exchange," against a *measure* of the entire market volume (*common* plus *preferred*) for the corresponding interval (Fig. 3). However, since the histograms giving the dispersion were taken for mid-month intervals, whereas published volumes were for month-end intervals, we have used as our measure of volume, the *arithmetic* average of the four or five weekly volumes most closely spanning

TABLE I

IDENTIFICATION OF STOCKS PROVIDING DATA OF FIG. 2

\$20–\$30 Class			\$80–\$120 Class		
Name	Initial price (approx.)	Median daily volume in round lots (approx.)	Name	Initial price (approx.)	Median daily volume in round lots (approx.)
Beech Aircr.	21	15	Ford	92	125
Am. Airlines	25	50	Nat. Lead	104	35
Sperry Rand	28	100	US Gypsum	90	15
Admiral El.	23	30	Dow Chem.	97	50
Minn. Moline	23	20	Merck	80	35
Dome Mines	20	25	Zenith Rad.	112	50
Dresser Ind.	29	30	Ow. Ill. Gl.	104	15
Am. Met. Cl.	25	50	Kennecott	95	35
Champl. Oil	20	15	Texaco	83	55
Sou. Pac. RR	23	60	Int. Paper	130	30
Freeport Sul.	25	30	Eastman Kod.	106	35
Celanese	27	50	Air Red.	82	15
	Average =	42		Average =	38

the time interval of the histogram. These volume averages were taken graphically from charts. Note that this procedure tends to remove a small *systematic* correlation between volume and dispersion because of the fact that not all mid-month intervals contribute the same number of trading days (*see* discussion of Table VI). A similar systematic error for the same reason presumably also affects the dispersion, which we have not attempted to correct. A possible corresponding correction to the dispersion would be to divide the observed dispersion by the square root of the number of trading or calendar days actually spanned for each mid-month interval. We have not applied such a correction, since it is too small to upset the conclusions drawn from Fig. 3.

The dashed line of Fig. 3 shows the theoretical dispersion developed for a market of stocks priced at \$40 at the beginning of each monthly interval, traded at an average rate of 1.5 round lots per transaction. Quantitatively it will be seen that the data are moderately well represented by the theory. The individual departures from the dashed lines are considerable, but in order of magnitude our simple picture is still maintained.

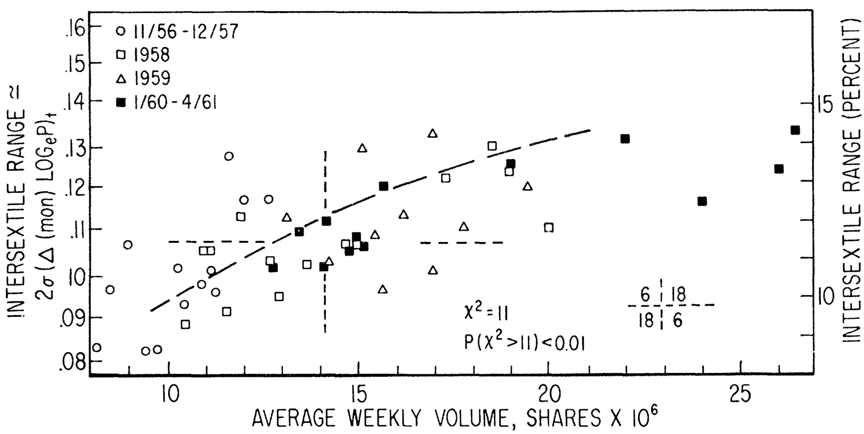

Fig. 3. Across-the-market dispersion (doubled) of $\Delta\log_e P$ for one month intervals, vs. volume for the same interval. Dispersion data were taken from histograms in "The Exchange" or "Monthly Review" published by NYSE. Volume data were taken graphically for the nearest coinciding 4–5 week interval from charts of Securities Research Corporation. The dashed line corresponds to the theoretical double dispersion developed in one month (4.3 weeks) by an ensemble of \$40 stocks diffusing in $\frac{1}{8}$ point steps, at the indicated volume rate in round lots$\div$1100, with 1.5 round lots per transaction$=$step. $2\sigma = 2[V_{(\text{wk, rd lots})} \cdot 4.3/1.5 \cdot 1100]^{1/2}(1/8.40)$. The sample size increased over the interval of the data 1957 to 1961 from 1075 to 1150. The number of shares listed on the entire Exchange increased from 4.5 to 6.5×10^9 over the same interval. The dashed line should be considered an approximation, since the sample size is $\sim$1100 common stocks, whereas the volume refers to the entire NYSE ($\sim$1600 issues), including preferred stocks.

In the interval 1956–61 covered by the data of Fig. 3, the total number of shares listed for trading increased by 33 per cent because of splits, stock dividends, and new listings. Presumably the average total market volume increased for this reason alone. It does not seem entirely obvious that this source of volume increase should increase the across-the-market dispersion of price changes according to the $\sqrt{V/1.5}$ rule, even though one purpose of splits and new listings is to increase trading. Therefore, in Fig. 4 we have plotted the data of Fig. 3, with the volume coordinate divided by the number of shares listed at the time the data were taken. This gives the abscissa as an 'average' fractional turnover per week.

Qualitatively, the same conclusion is reached as before. When the market activity, or turnover, is large, so also is the across-the-market dispersion. We can no longer plot a theoretical dashed line (as in Fig. 3) for a hypothetical \$40 stock, since we do not have a theory relating turnover to the price dispersion developed.

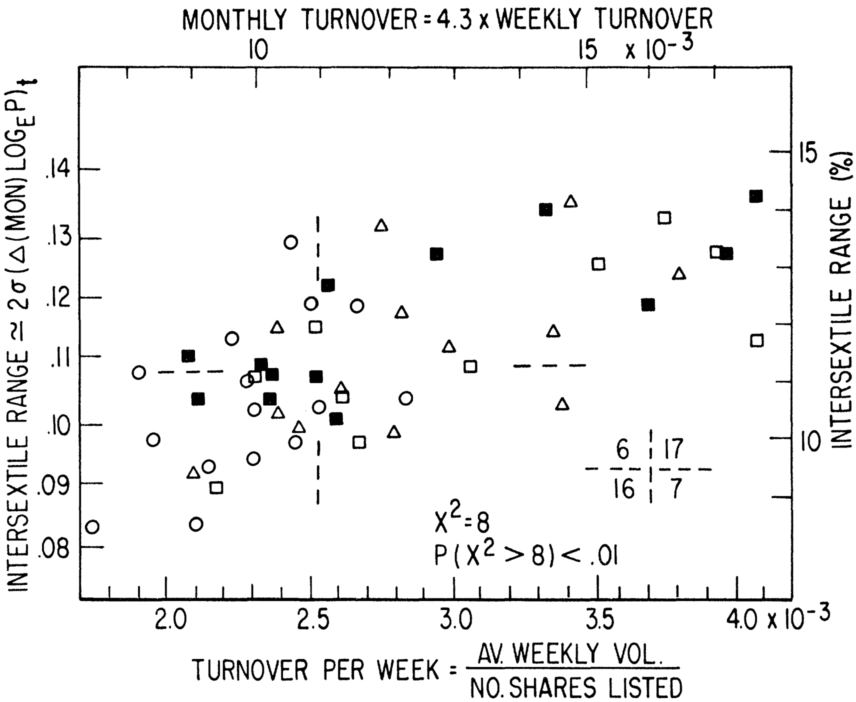

Fig. 4. Across-the-market dispersion (doubled) of $\Delta \log_e P$ vs. 'average' turnover. The data are from Fig. 3, with volume data divided by the total number of shares listed for each interval; data taken from the "Monthly Review." The numbers in the corners of Figs. 3 and 4 are contingency tables about the median, and P is the probability of χ^2 being greater than the observed value of χ^2 for no statistical dependence.

THE VOLUME SEQUENCE

In view of our only partial success in relating the dispersions of the price change sequences to the volume sequences, it might be interesting to try to answer the simple question—What are the properties of the volume sequences? We have a theory that describes, at least approximately, how the price sequences behave; it is equally interesting (at least from a purely scientific standpoint) to try to make a theory for the volumes.

According to Fig. 1, the vast majority of individual transactions are for 1–2 round lots, and according to references 1 and 2, it is sufficient to assume that successive changes of $\log_e P$ are approximately independent in the probability sense. Given some degree of success of this assumption for the price changes, let us make similar assumptions concerning the volume. Accordingly, we assume (A) that the probability of one transaction ($\cong$1.5 round lots) occurring in any finite span of time τ, is independent of the occurrence of any other transaction in that same span. We further assume (B) that the probability of occurrence is proportional to the length of τ. If these assumptions, which are both plausible and justified for many physical situations, are correct also for the occurrence of a stock market transaction, then we can make the following predictions about the volume and its fluctuation, or sequential dispersion for any interval of time length τ.

1. If $V_j(\tau)$ is, say, the daily ($\tau = 1$ day) volume of the jth stock, then the number of transactions $T_j = V_j/1.5$ should follow a Poisson distribution (with j fixed). If $\bar{T}$ is large (and in practice 10 or more would be large enough) then T should be approximately *normally* distributed with a sequential dispersion, $\sigma_T = \sqrt{\bar{T}}$, or for the volume $\sigma_V = \sqrt{1.5\bar{V}}$. Numerically, then a stock for which the arithmetic mean daily volume was 100, the individual daily volumes would be normally distributed about 100, with a sequential dispersion $\sigma_V = \sqrt{150} = \pm 12.2$.

2. As a percentage of mean daily volume, the sequential dispersion of daily volume for different stocks would decrease as the square root of the mean daily volume. Equivalently, the Poisson law predicts $\sigma[\log\ V(\text{day})]_j \cong \sigma[V(\text{day})]_j/\bar{V}(\text{day})_j) = \sqrt{1.5}/\sqrt{\bar{V}(\text{day})_j}$.

3. If one compares the sequential dispersion of volume for a single stock on a daily basis with volumes for the *same stock* on a weekly or monthly basis, then the percentage fluctuation or dispersion should decrease as the square root of the respective time intervals. Thus $\sigma[\log\ V(\tau = \text{day, week, month})]$ should be in the ratios 1, $1/\sqrt{5}$, $1/\sqrt{22}$ where the numbers are the number of trading days in the specified intervals. Since τ and $\bar{V}(\tau)$ should be proportional to each other (assumption B), we may equivalently say that $\sigma[\log\ V(\tau)]_j \cong \sigma V(\tau)_j/[\overline{V(\tau)}]_j = \sqrt{1.5}/\sqrt{\overline{V(\tau)}_j}$.

One should note the distinction between **2** and **3** above. In **2** we increase the volume by passing to different stocks (Fig. 6), keeping the time interval fixed. In **3** we stick to the same stock, and increase the volume by increasing the interval (Fig. 7).

Comparison of these conclusions from theory with observation reveals some rather gross discrepancies, though the theory does give some qualitatively correct predictions.

Figure 5 tests, for a number of different stocks and the entire market, the first conclusion that the sequential distribution of daily volumes of individual stocks should be normal. Generally speaking, the distribution is more nearly log normal, 5(b), than normal, 5(a), contrary to our prediction. The plots under (*b*) are more nearly straight than those for (*a*). For the largest daily volumes, Standard Oil of New Jersey and the entire market volume, the hypotheses of normality or log-normality fit the data just about equally well, as indeed is to be expected for either a true normal

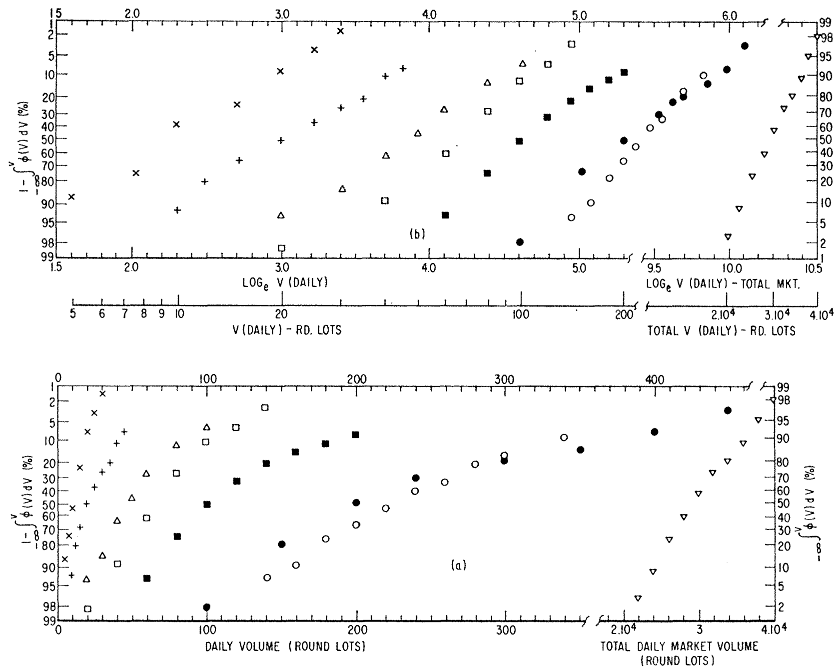

Fig. 5. Cumulated sequential frequency distributions of daily volumes of individual stocks, and entire market (broken scale, right) plotted to test for normal (*a*) vs. lognormal (*b*) distribution. *X* Pullman, June–Dec. 1958; + Sou. Cal. Ed., June–Dec. 1958; △ Phelps Dodge, Oct. 1957–May 1958; □ Texaco, Oct. 1957–May 1958; ■ Am. Cyanamid, June–Dec. 1958; ● U. S. Steel, June 1958–Jan. 1959; ○ Stan. Oil (N.J.), Oct. 1957–Apr. 1958; ▽ total NYSE market volume, Jan.–Dec. 1960. The average sample size for 6 months of data indicated is about 130.

or log normal distribution whenever the ratio dispersion/mean becomes considerably smaller than unity.

Figure 6, using the data of Fig. 5, plus a few additional samples of data, tests the second conclusion. Figure 6 uses the median volume as a measure of the mean, a fair approximation except when $\sigma(V)/V_{\mathrm{med}} \cong 1$. If one compares the dispersion indicated on Fig. 6 with the theoretical values according to the Poisson distribution, the observed values are seen to be much too large. Figure 6 also shows that the sequential dispersion of $\log_e V(\text{day})$ does decrease with $V(\text{day})$, but much less rapidly than $1/\sqrt{\text{Volume}}$ as the theory predicts.

Figure 6 shows the sequential dispersion of log V(day) as a function of increasing average V(day) in passing *from one stock to another*. Figure 7 [to test conclusion (3)] shows the same effect when the length of time-interval determining the volume is increased for *the same stock*. The data of Figs. 6 and 7 tend to agree with each other, as they should if assumptions A and B were correct. Quantitatively, Figs. 6 and 7 both disagree with the theory.

It would be quite possible from a purely artificial viewpoint to devise a model in which successive *daily* volumes were *independently* and *log normally* distributed. If this were true of the data, then by central limit

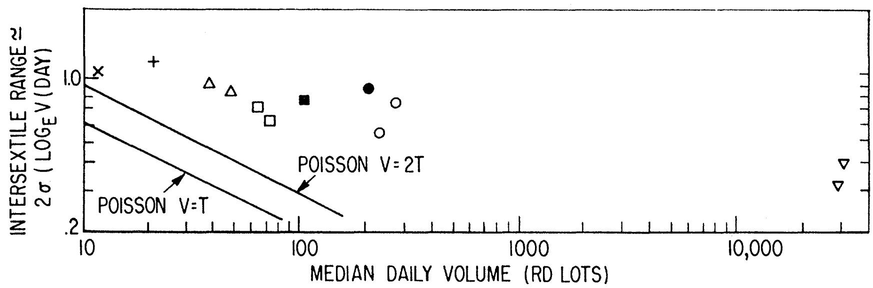

Fig. 6. The sequential dispersion (doubled) of log daily volume in round lots vs. median volume, both taken from Fig. 5, plus a few additional samples of data (1957–60) for the same stocks, and entire market. The straight lines give the theoretical value of the double dispersion according to Poisson's law for the observed volumes, under the assumption of 1, or 2 round lots per trade. That is, $2\sigma(\log V)=2/\sqrt{V}$, or $=2\sqrt{2/V}$.

theorem the weekly and monthly volumes would approach *normality*, with sequential dispersions of $\log_e V(\tau)$ decreasing as we mentioned, as $1/\sqrt{V}$. In fact this is not so. Figure 7 also shows the decrease of $\sigma[\log V(\tau)]_j$ is much less rapid than $1/\sqrt{V}$.

It should be noted that the data of Fig. 7 come from samples of 12 (day, week, month) *consecutive* values of volume. Were the 12 values nonconsecutive, and synthetic weeks or months composed of nonconsecutive daily data, one might expect the inverse square root dependence on V for $\sigma[\log V(\tau)]_j$ to be more nearly fulfilled. We have not checked this conclusion directly.

Figure 8 gives two examples of the details of the sequential distribution of monthly volume. Figure 8 shows one example (a) which actually turned out to be approximately normal, and one (b) approximately log normal (both stocks also appear on Fig. 5, where their daily volumes are more nearly log normal). Taken together, Figs. 7 and 8 indicate that in general a month (22 trading days) is not long enough for the central limit theorem to normalize the sum of 22 single day's trading.

The conclusion that we must draw from the observed size of the sequential dispersion of the volume, relative to the Poisson law predictions (Figs. 6 and 7), is that some intervals of time have a much greater probability to contain transactions than others of equal length, contrary to our original and physically plausible assumptions (*A*) and (*B*). The above

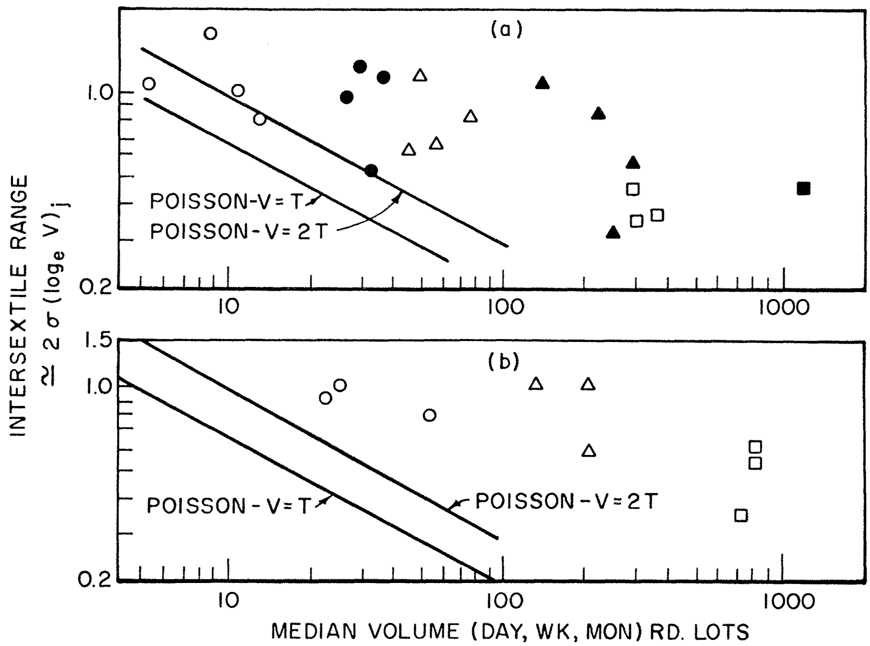

Fig. 7. Sequential dispersion (doubled) of daily (○), weekly (△), monthly (□) log volumes, vs. median volume of the same data. Dispersion and median were determined from twelve *consecutive* values of the daily, weekly, or monthly volume. (*a*) Firestone: filled data points refer to after the 3 for 1 split on March 24, 1960; open points, prior to the split. (*b*) United Airlines. Daily data samples for (*a*) and (*b*) taken in 1960–61, weekly in 1957–61, monthly in 1956–60. The lines give the theoretical curve according to Poisson's law (*see* legend Fig. 6).

conclusion, laboriously extracted from the nonagreement of Poisson's theory with observations, is a statement of a well-known property of the market—a tendency for stock to be traded in concentrated bursts. In fact, it is one of the major problems of governing the Exchange, to prevent the concentrated bursts of orders from completely disrupting trading.

Figures 7 and 8b give some indication of the duration in time of the bursts. According to these figures, a month is not long enough to smooth out the nonrandom occurrence of bursts, and satisfy assumption (*B*) so that the percentage fluctuation in volume, or $\sigma[\log V(\tau)]_j$ would decrease like $1/\sqrt{V}$ or $1/\sqrt{\tau}$. In other words, stocks suffer extended periods of excessive 'interest' (cf. volume), or lack of interest, lasting a month or more. This conclusion seems to be correct, and actually a quite conservative statement about the duration of fads, or their opposite, aversions, in the stock market.

The fact that the daily volume tends to be log normally, rather than

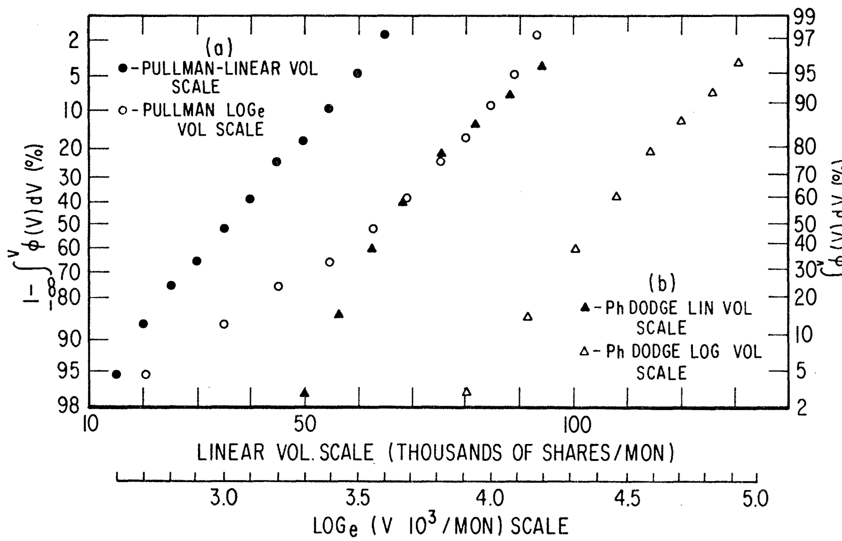

Fig. 8. Cumulated sequential distribution of monthly volume, tested for normal (filled data points) vs. lognormal (open points) distribution. (*a*) Pullman, data from June 1945 to June 1957, $N=144$; (*b*) Phelps Dodge, data from June 1952 to June 1961, $N=108$. Data taken graphically from charts of Securities Research Corporation. No splits occurred in the span of data.

normally distributed, can be expressed 'theoretically' in a plausible way by saying that: if volume measures 'interest' or 'attention' to a stock, then the increment (random) of interest is proportional to the interest already present; i.e., people, like sheep, tend to develop more interest because it is already there (and conversely). This tendency of people is well-known to professional manipulators, who attempt to generate real interest by producing a semblance of interest by spurious trading. This statement of a possible 'derivation' of the log normal sequential distribution of daily volume, is one way of expressing the failure of Poisson's law. Evidently, if the presence of large volume 'induces' bigger volume fluctuations, then the condition of independence of individual trades [assumption (A)] necessary to derive Poisson's law is not fulfilled.

The method of analysis yielding the above conclusions about the nonuniform rate of trading in the stock market was drawn from two physical analogs of interest. Let us first suppose that the stock ticker is in fact a cosmic ray Geiger counter, which records the 'ionization' (separation of buyer from his money, and seller from his stock) associated with the passage of a particular elementary particle (name of stock) through the 'ionization chamber' (floor of the Exchange). By studying the counts on our instrument we can decide if there are bursts of unusually frequent occurrence for any particular type of elementary particle, and how long these bursts last. We conclude that the counts for our particles, as for the count on real Geiger counters, are not uniformly distributed in time even when we average or sum over intervals as long as a month.

A second analog concerns the distribution of the stars in space. Here, the star counts on stellar photographs compared with the prediction of the Poisson law show, unambiguously, that stars are nonuniformly distributed in clusters in space, the clusters themselves are nonuniformly clumped into star clouds or galaxies, which are themselves in turn clustered much more than a random distribution of galaxies would indicate. Similarly, stock trading is clustered into days of unusual activity, the days into weeks, the weeks into months and months into seasons of excessive activity, much more than a constant probability of occurrence in time would allow.

PERIODICITY IN TIME

IN VIEW OF the preceding discussion, it does seem reasonable to suppose that there is a statistical relation between the total market volume developed over a day, week, or month, and the across-the-market dispersion in price changes developed over the same interval. Figures 3 and 4 showed this directly for intervals of a month; it seems reasonable to suppose a similar effect for intervals of a day, a week, or even a year. In what follows, we hope to show that there is definite evidence of strictly periodic behavior in the volume for intervals of a day, week, three months, and one year, and hence by implication a periodic oscillation in the across-the-market dispersion of price change for the same interval of time. In this way we will infer that there is a time periodic structure in price. It will be noted that this periodicity is not of a type that would show up either in Fourier or correlogram analysis of a single stock price sequence, or of a single index sequence, nor can it be exploited for profit in any simple way. This periodic dispersion is a periodic property of the ensemble.

Figure 9 shows a plot of the hourly rate of trading volume for several successive days. The data show, as is in fact well-known, that there is a quite characteristic and reproducible burst of trading at the beginning and end of the trading day. Note that the 18½-hour gap between trading sessions has been considerably compressed in Fig. 9. The question we now ask is: How does the across-the-market dispersion of price changes during the $5\frac{1}{2}$ hours of trading compare with that for the $18\frac{1}{2}$ hour interval of no trading? We can make this comparison by computing the across-the-

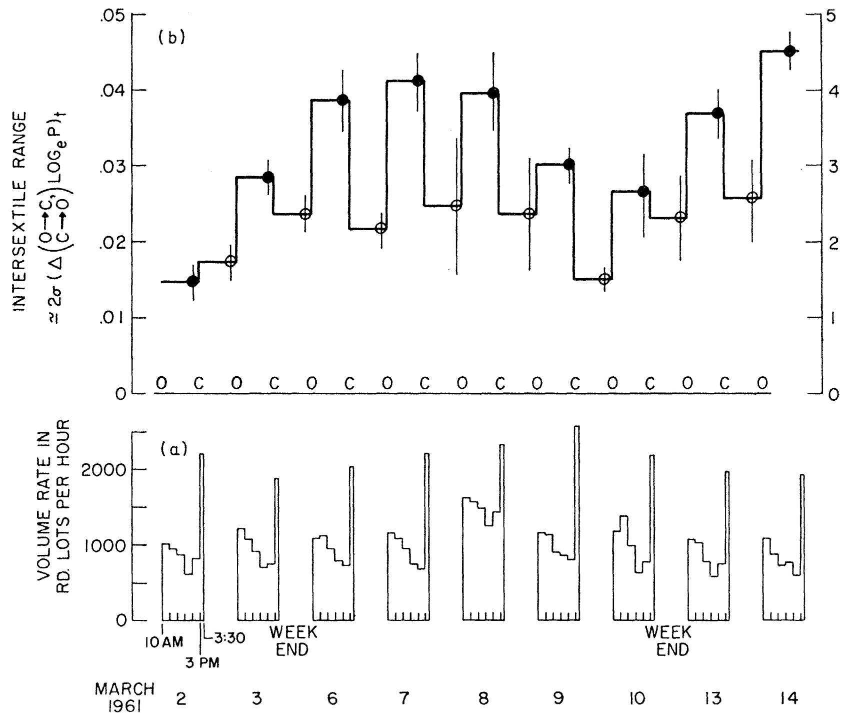

Fig. 9. (*a*) Hourly volume rate of trading vs. calendar time, NYSE. The trading of the last half-hour (3–3:30 P.M.) has been doubled to give the hourly rate. (*b*) The across-the-market dispersion (doubled) of 'open' to 'close' (filled circles) and 'close' to 'open' (open circles) changes in $\log_e P$ for a sample of 12 NYSE common stocks, taken for the same time interval as (*a*). The sample was taken under the letter N—Nafi to Nat. Fuel Gas; data from *Wall Street Journal*. The range of uncertainty (light line) corresponds to the largest range from interval, 2nd to 3rd, or 10th to 11th member of the sample.

market dispersion of open to close price changes, and the same dispersion of close to open prices, all of which are published in the *Wall Street Journal*. In this way we can compare the dispersion for mostly volume with small passage of time ($5\frac{1}{2}$ hours) with mostly time ($18\frac{1}{2}$ hours) and little volume. In practice the volume from 'close' to 'open' prices contains the opening block.

Figure 9b shows that the price dispersion oscillates in phase with the volume, and gives a comparison of the across-the-market dispersion computed from close to open price changes, with open to close price changes. It will be noticed that the latter dispersions are appreciably larger than the former, but much less than if the dispersion were strictly proportional to $\sqrt{\text{volume}}$, since we have pointed out, the volume from 'close' to 'open'

TABLE II

(*a*)

(The number of days of increase or decrease of daily volume of odd lot buying, *VOLB* over previous day, for different days of the week. Holidays and post-holidays excluded from the count.)

	Mon.	Tues.	Wed.	Thurs.	Fri.	
$\Delta VOLB>0$	44	12	15	22	19	*Data from 1/2/58 to 12/31/58*
$\Delta VOLB<0$	3	38	36	27	28	
$\Delta VOLB>0$	43	15	17	22	20	*Data from 1/7/59 to 12/31/59*
$\Delta VOLB<0$	4	34	34	29	27	
$\Delta VOLB>0$	37	11	21	24	20	*Data from 1/4/60 to 12/1/60*
$\Delta VOLB<0$	5	32	26	23	25	

(*b*)

(Contingency table test of the above data for the dependence of daily odd lot buying volume 'increasing' or 'decreasing' on 'day of the week.')

	Mon.	Tues.	Wed.	Thurs.	Fri.	Total
$\Delta VOLB>0$	124	38	53	68	59	342
$\Delta VOLB<0$	12	104	96	79	80	371
Total	136	142	149	147	139	713

$\chi^2=138$; $P(\chi^2>138)\ll 0.01$; $n=4$ degrees of freedom

is the opening block, usually only a small fraction of the total day's trading volume. Figure 9b confirms, on a daily basis, what we had already inferred from the monthly data of Figs. 3 and 4. Hence, we can conclude from the data that the across-the-market dispersion of stock price changes has a rather well-defined diurnal period, when prices are considered as a function of calendar, or total elapsed time, rather than a function of simple trading time interval, for which 1 day $=5\frac{1}{2}$ hours. These results have the quite plausible and obvious interpretation that volume represents interest or attention to stocks, and that prices tend to move under the impact of this interest. Since the sign (up or down) of the motion is unspecified, the effect of the 'attention' shows up in the across-the-market dispersion, which fluctuates periodically in the diurnal cycle of man's attention span.

The next question is: Can one detect periods of greater length? The diurnal cycle is obvious in the data, but the demonstration of the presence of longer periods will require a somewhat more subtle statistical analysis. Casual examination of the daily volume of either individual stocks, or for the entire market, does not reveal any structure as obvious as that in the hourly volume shown in Fig. 9a. Nevertheless, there are certain small aspects of the market volume which do repeat themselves rather regularly on a weekly basis. Table II gives by day the number of occasions the odd lot buying increased or decreased over the preceding day (on Monday the preceding day would be Friday). Odd lot buying volume alone amounts to about 10 per cent of the total market volume, as well as the odd lot selling volume. It should be noted that the published figures for total market volume refer to the sum of round lots only, exclusive of odd lots. The relation of odd to round lot market is fully discussed in reference 5.

It will be seen that there is a quite systematic tendency for odd lot buying to be larger on Monday than on Friday and to subside throughout the remainder of the week. This structure is of course well-known, and is sufficiently marked to be obvious to the eye when the odd lot buying or selling data are plotted sequentially. We only give these data in order to make comparison with other effects described below. A χ^2 test of the summary 2×5 contingency table, containing the number of occasions in which the volume of odd lot buying increased or decreased from the previous day, vs. day of the week, shows that it is most unlikely that the average probability of an increase or decrease (as given by marginal row totals) is in fact independent of the day of the week. Hence, we conclude that the odd lot buying has a significant weekly periodicity.

This periodicity has the simple and obvious interpretation that the odd lot trader, professionally engaged in other matters during the week, tends to make up his mind over the week-end when he has the time to do so, and acts on Monday. The distribution in time of information from professional market advisors is geared to this schedule, as this information also tends to come over the week-end. However, this distribution in time deserves closer study, in view of the different type of periodicity of the total market volume discussed below.

If we examine the total market volume (Table III) in a manner similar to that in which the odd lot buying was examined, we see from the value of χ^2 and the associated probability that there is also a significant weekly periodicity in the total market volume. It is a great deal less pronounced than the periodicity in the odd lot buying, as it is just barely significant at the 1 per cent level. More interesting, the significant increases for total market volume occur on Tuesday and Wednesday, whereas for odd lot buying it was most pronounced on Monday. From this weekly periodicity in the volume we infer a weekly periodicity in the daily across-the-market dispersion of stock price changes. Presumably this price dispersion is a maximum on Tuesday or Wednesday. One can interpret the data as meaning that round lot traders, in the market as a business, tend to forget

TABLE III

(*a*)

(The number of days of increase or decrease of total daily market volume, V, over previous day, for different days of the week. Holidays and post-holidays excluded from the count.)

	Mon.	Tues.	Wed.	Thurs.	Fri.	
$\Delta V>0$	21	29	29	25	21	*Data from 1/2/58 to 12/31/58*
$\Delta V<0$	26	21	22	24	26	
$\Delta V>0$	24	30	28	21	18	*Data from 1/2/59 to 12/31/59*
$\Delta V<0$	23	19	23	30	29	
$\Delta V>0$	19	24	28	21	18	*Data from 1/4/60 to 12/2/60*
$\Delta V<0$	23	19	19	26	28	

(*b*)

(Contingency table test of the above data for the dependence of volume 'increasing' or 'decreasing' on 'days of the week.')

	Mon.	Tues.	Wed.	Thurs.	Fri.	Total
$\Delta V>0$	64	83	85	67	57	356
$\Delta V<0$	72	59	64	80	83	358
Total	136	142	149	147	140	714

$\chi^2=13.5$; $P(\chi^2>13.5)<0.01$; $n=4$ degrees of freedom

it over a long week-end, make up their minds on Monday and Tuesday, and act the next day.

In addition to the daily data on prices and volumes, there is also published on each date t, four sequences of numbers that can be examined as indicators of the state of the market. These are $N_a(t)$, $N_d(t)$, $N_u(t)$, and $N_I(t)$, which are respectively the number of prices which advanced, the number which declined, the number unchanged (all from previous day or the last day traded), and the sum of these three, the number of issues traded. The last one, of course, fluctuates much less percentagewise, than any of the other three. The running sum $B(T)=\sum_{t=t_0}^{t=T}[N_a(t)-N_d(t)]$

is the so-called breadth of market index. $B(T)$ represents approximately the median of prices, and by adjusting the scale of its plotting it can usually be made to be approximately parallel to the well-known arithmetic averages that represent the mean. For simplicity's sake, let us consider just the sequence $N_a(t)$, and examine it in a manner identical to that which was used to show weekly periodicity in the daily total and the odd lot buying

TABLE IV

(The number of days of increase or decrease in the 'number of stocks advancing,' N_a over previous day, for different days of the week. Holidays and post-holidays excluded from the count.)

	Mon.	Tues.	Wed.	Thurs.	Fri.	
$\Delta N_a > 0$	20	22	26	30	24	*Data from 1/2/58 to 12/31/58*
$\Delta N_a < 0$	27	28	25	19	23	
$\Delta N_a > 0$	12	24	26	22	28	*Data from 1/2/59 to 12/31/59*
$\Delta N_a < 0$	35	25	75	29	19	
$\Delta N_a > 0$	7	26	14	26	27	*Data from 1/4/60 to 12/2/60*
$\Delta N_a < 0$	35	17	28	21	19	

(Contingency table test of the above data for the dependence of N_a 'increasing' or 'decreasing' on 'day of the week.')

	Mon.	Tues.	Wed.	Thurs.	Fri.	Total
$\Delta N_a > 0$	39	72	71	78	79	339
$\Delta N_a < 0$	97	70	78	69	61	375
Total	136	142	149	147	140	714

$\chi^2 = 28.0$; $P(\chi^2 > 28) < 0.01$; $n = 4$ degrees of freedom

volume. The data are summarized in Table IV, and it is seen that with a chance of error less than 1 per cent we can conclude there is periodicity in $N_a(t)$. Moreover, it can be easily shown that most of the contribution to χ^2 occurs on Monday, when the odds are approximately $2\frac{1}{2}$ to 1 that $N_a(\text{Monday}) < N_a(\text{Friday})$.

This decrease of N_a from Friday to Monday does *not* mean that 'the market' (as measured by the breadth of market index) tends to drop on Monday relative to Friday, or that the 'velocity' (first difference) of prices is downward over the week-end. The decrease in N_a does mean that either (1) the 'acceleration' of the median price changes, or second difference is negative over the interval Thursday–Friday–Monday; this also says that the breadth of market index tends to be concave down over Thursday to Monday. (2) A second possible interpretation is that the total number of issues traded (and perforce N_a also) tends to drop from Friday to Monday.

This second alternative is eliminated when we examine Table V where simultaneous data on N_a and N_I are assembled. Here it will be seen that the number of issues N_I tends to *increase* systematically on Monday relative to Friday, so that the Monday decline of N_a is to be interpreted as a genuine tendency for negative or downward *acceleration* of prices over the

TABLE V

(Contingency table test for the number of advances, N_a, and number of issues traded, N_I, increasing or decreasing, vs. day of the week, both for the interval 4/8/60 to 6/23/61. Data from *Barron's Weekly*. A few weeks' issues were missing from the file, which accounts for smaller number of Monday (Friday to Monday changes) data points.)

	(a)					
	Mon.	Tues.	Wed.	Thurs.	Fri.	Total
$\Delta N_a > 0$	17	27	27	28	36	135
$\Delta N_a < 0$	33	27	28	29	21	138
Total	50	54	55	57	57	273
	$\chi^2 = 8.5$; $P(\chi^2 > 8.5) = 0.07$; $n = 4$ degrees of freedom					
	(b)					
$\Delta N_I > 0$	35	27	21	22	25	130
$\Delta N_I < 0$	15	27	34	35	32	143
Total	50	54	55	57	57	273
	$\chi^2 = 13.6$; $P(\chi^2 > 13.6) < 0.01$; $n = 4$ degrees of freedom					

interval Thursday, Friday, Monday, even though prices may be rising over the same interval.

One can compare the successive values of monthly volume to determine the presence of a significant annual periodicity therein, in a manner identical to that in which daily volumes were examined to show periodicity of one week. This is done in Table VI(a), which shows that it is most unlikely that these monthly fluctuations could have occurred by chance, under the hypothesis of no periodicity. However, certainly a part of the large value of χ^2 must be a consequence of the fact that the months are of unequal length in days; i.e., the fact that the February volume was less than January volume on thirty-three out of forty-five occasions must by in part because of the fact that there are fewer days in February on which to trade. We have endeavored to remove this contribution to χ^2 by appleing correction factors to the individual monthly volumes, which are listed in Table VI(b). For example, 90 per cent of the January volume was

TABLE VI

(Contingency table tests of total monthly volume of *NYSE* increasing or decreasing from previous month, vs. month of the year. Data from 1915–1960.)

(*a*) *Uncorrected*

	J	F	M	A	M	J	J	A	S	O	N	D	Total
$N(\Delta V>0)$	16	12	35	16	20	20	20	21	33	32	23	29	276
$N(\Delta V<0)$	29	33	10	29	25	25	25	24	12	13	22	16	264
Total	45	45	45	45	45	45	45	45	45	45	45	45	540

$\chi^2=55$; $P(\chi^2>55)\ll 0.01$; No. of d.f. $=11$

(*b*)
Correction factors to monthly volume, based on number of days in month, less holidays

	J	F	M	A	M	J	J	A	S	O	N	D
No. of days N_j	30	27	31	29	30	30	30	31	29	31	28	30
correction factor N_{j+1}/N_j	0.9	1.15	0.935	10.35	1.0	1.0	10.32	0.935	10.68	0.902	1.072	1.00

(*c*) *Corrected*

	J	F	M	A	M	J	J	A	S	O	N	D	Total
$\Delta V>0$	16	18	25	22	18	20	20	16	35	28	26	22	266
$\Delta V<0$	29	27	20	23	27	25	25	29	10	17	19	23	274
Total	45	45	45	45	45	45	45	45	45	45	45	45	540

$\chi^2=30.5$; $P(\chi^2>30.5)<0.01$; No. of d.f. $=11$

compared with February volume for increase or decrease of volume and the results tabulated in Table VI(c). The value of χ^2 is almost cut in half, but even so we can still conclude to a significant annual periodicity in our monthly volumes.

The correction factors in Table VI(b) were determined by successive ratios of the number of calendar days diminished by the number of holidays. One might have preferred to use the actual number of trading days in each month of each year, though it is not entirely clear that a correction based only on trading time interval rather than calendar interval is necessarily required. The actual corrections chosen were a compromise based primarily on simplicity and ease of application. The volume data were taken graphically from monthly charts of the Securities Research Corporation, so that the correction factor only needed to be applied when visual inspection of the data suggested that use of the factor would change a volume increase between successive months to a decrease, or conversely. The proper correction to apply when examining monthly data with a computer is one which might well be examined more closely.

Market reporters and statisticians regularly summarize the record for months in which the market (as measured by an arithmetic average) advanced or declined, and there are well-established traditions of the market place that certain months are more likely than not to advance (i.e., the summer rally, or decline in winter months). We can examine this record of months of these advances and declines by the same contingency table test used previously.

Our particular source of data is GRANVILLE,[6] who gives the percentage of years in which each month showed a rise or fall of the Dow Index for the years from 1886 to 1960. Accepting his percentages as correct, we can convert them to frequencies of occurrence and test for the significance of an annual periodicity; colloquially, the genuineness of the summer rally, or winter as the season of discontent.[7]

It will be seen (Table VII) that the probability that the observed variations of advances and declines in each month could have occurred under the condition of identical probabilities for each month is 0.011. Thus the tradition of the market place is in fact supported by the evidence, but only barely so. A '1 per cent significance' statistician would reject the hypothesis of a real periodicity, whereas a '5 per center' would accept it. Personally we feel the effect is genuine, but small.

However, it should be noted that the above test, accepted as giving a significant result, does not say that the *mean* of the market index changes is significantly positive or negative in certain months. It does make this statement about the median. The mean, and hence expectations, may indeed also be significantly positive or negative in certain months, but strictly speaking, the test does not tell us this.

The inference from the weekly or annual periodicities in the volume is that the across-the-market dispersion of price changes would also show the same periodicity, since volume and dispersion were statistically connected. We give in Fig. 10 and Table VIII some additional indirect evidence that volume, hence also the across-the-market dispersion of price changes, should have a three-month and also an annual periodicity. It is commonly believed that there is some statistical relation between the earnings sequence, earnings changes, the price sequence, and price changes. Without going into the question of what the statistical relations between these four sequences are, or when they are most pronounced, it would appear reasonable to suppose that if there is some relation between prices and earnings, the dates on which the information about earnings becomes available would be related to the price changes at about the same time

TABLE VII

(Contingency table test for advance and decline of market average in a month, for different months of the year. Data from 1886–1960 from reference 5.)

	J	F	M	A	M	J	J	A	S	O	N	D	Total
N(market advance)	38	26	26	38	40	40	44	47	30	36	36	37	438
N(market decline)	36	48	48	36	34	34	30	27	44	38	38	37	450
Total	74	74	74	74	74	74	74	74	74	74	74	74	888

$\chi^2 = 24.6$; $P(\chi^2 > 24.6) = 0.011$; $n = 11$ degrees of freedom

Now information on earnings is not distributed uniformly throughout the year, but is distributed in bursts about three months apart, with a sustained burst of information after the end of the year, corresponding to those companies that report annually (*see* Fig. 10).

Not knowing whether the information contained in these reports is good or bad, or the extent to which the information has already been discounted, it would seem reasonable to suppose that concentrated earnings information would tend to increase, relatively, the across-the-market dispersion of price changes at the time when information becomes available. Granville[6] assesses the effect of the concentration of earnings reports on trading procedure.

MERJOS[8] gives data that show the tendency of prices to respond to the prospect of presumably favorable information. A χ^2 test of the 2×2 contingency Table VIII 'before' and 'after' news, vs. price changes greater or less than zero, shows unambiguously that even the prospect of presumed favorable information, without knowing what the information might be, makes prices jump in a systematic fashion. The Table shows that the prices tend to rise before the supposed good news, and fall afterwards, so that dispersion is very likely to be largest on the day the most news is released.

If we compare the above conclusions on periodicity with investigations in the past searching for periodicities in stock prices, it will be seen that the search has been primarily for periodicities in the prices themselves, or in the first differences of prices. It is the first moment of these price changes that measures the expectation of profit or loss.

It might be well to comment here on the interpretation of our contingency table tests as indicating periodicity, with the results of a correlogram or Fourier analysis. A significant χ^2 test says that the fivefold attribute 'day of the week' is significantly statistically related to the twofold attribute; for example, volume 'increasing' or 'decreasing' from the previous day. If our interpretation of periodicity is correct, it should mean that the fifth serial correlation coefficient of daily volume changes is positive and significantly different from zero (alternatively, that the Fourier coefficient for a 5-day period is significantly different from its value for a purely random sequence of numbers).

TABLE VIII

(Contingency table test for the effect on stock prices of announcing a speaker giving 'news' at 129 lunch meetings of New York Society of Security Analysts. Zero changes omitted from data.)

Price change	Interval of 2 weeks before news	Interval of 2 weeks after news	Total
$\Delta P>0$	76	51	127
$\Delta P<0$	44	73	117
Total	120	124 $\chi^2=12$; $P(\chi^2>12)\ll 0.01$	244

We have not been able either to prove or disprove the equivalence of our χ^2 test to a significance test applied to a correlogram or periodicity analysis. In order to ensure that our positive results were not some subtle statistical artifact, such as might be associated with the known first serial correlation of first differences in a simple sequence of independent random numbers, we have evaluated the contingency table 'increase or decrease' of successive random numbers, using standard tables of random numbers. Such tables are grouped in units of five numbers, conveniently analogous to data from a 5-day week. No significant periodicity from our χ^2 test was found in this case. Here we may safely conclude that our positive results for a week, or a year, are not the result of a statistical artifact.

PERIODICITY IN THE PRICE COORDINATE

LET US NOW consider a different type of periodic structure, in the price coordinate.[9] The notion of a periodic space structure is a familiar one in physics, which is replete with examples of 1-, 2-, or 3-dimensional periodic space lattices. In the stock market the price or space coordinate structure, which we are about to describe, is conventionally called the phenomenon of resistance and support,[10] or in economics the TAUSSIG[11] penumbra. There are actually two aspects of this phenomenon, which can be separately interpreted as the aspect of congestion and the aspect of reflection.

The congestion aspect states that there are price ranges in which a given stock price spends an inordinate amount of time; equivalently, there

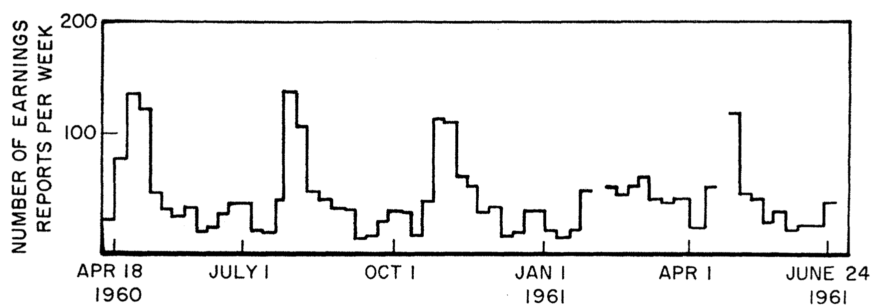

Fig. 10. Number of earnings reports per week vs. calendar time. Data from *Barron's Weekly*, under letter A through F, NYSE. Gaps simply correspond to missing issues from the file.

are many more transactions in this range than would be expected by chance. The reflection aspect states that there are price levels at which a stock is more likely than could be expected by chance to be reflected, or turned back in a direction opposite to its previous motion, up or down. If the particular reflection 'barrier' is above the price, the barrier is called resistance; if below it, support. It is commonly believed that at least some resistance and support levels occur together; i.e., once a resistance barrier has been crossed, this same barrier will provide support on a future decline. The above description, for us, is that of Brownian motion in the presence of partly reflecting barriers. They need not have the same reflectivity from both sides, and the reflectivity may vary with time.

The question that we ask here is, is it possible to show that the above described phenomena actually exist, in a statistically significant sense? When we consider simultaneously the ensemble of sequences of stock prices, it is then possible to show that both the aspects of congestion and reflection do exist, in the sense that we have described them. To show the congestion aspect, we have plotted in Fig. 11 the distribution of closing prices, using only the last eighth. It will be seen that there is a pronounced tendency for prices to cluster on whole numbers, halves, quarters, and odd one-eighths in descending preference, like the markings on a ruler. The nature of the effect may be readily verified by counting in any issue of the newspaper, the relative frequency of closing prices which are even or odd eighths. It should also be noted by comparing Figs. 11(a) and (b) that the effect is more pronounced with a sample from the entire market than if it is a sample restricted to a weekly volume greater than 50 round lots.

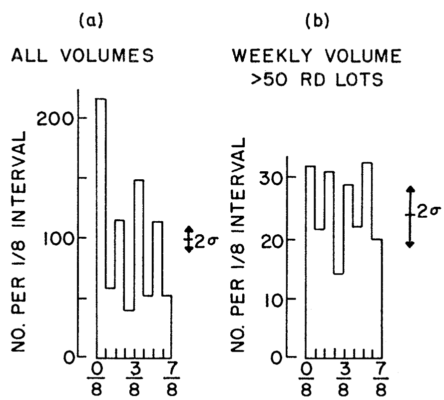

Fig. 11. Distribution of closing prices in eighths on NYSE, January 15, 1960, common and preferred. (*a*) Sample taken under A–K, all volumes. (*b*) Sample taken under A–E, for weekly volumes greater than 50 round lots. Data from *Barron's Weekly*. The double arrow denotes the mean and expected ±dispersion, or fluctuation of the number in a class, for a uniform distribution. $\sigma = \sqrt{N/8}$, where N is the total sample number.

This means that the ruler, or congestion effect, is more pronounced the smaller the volume.

The 'partially-reflecting-barrier' aspect of resistance and support can be demonstrated in the following way. If there are partially reflecting barriers in the one-dimensional 'field' in which prices move, then maxima and minima will tend to cluster on these barriers, more so than might be expected for random walks without such barriers.

Now the published record contains the closing price, as well as the high or low for a preceding fixed period (day, week, month, or year). This high or low is a maximum or minimum for the data in the fixed preceding interval only. But a 'high' for example, may or may not be a maximum, if a 'maximum' is defined as a price greater than the nearest *different preceding* or *following* price. A low may or may not be a minimum, if minimum is similarly defined. In order to 'enrich' the content of 'highs'

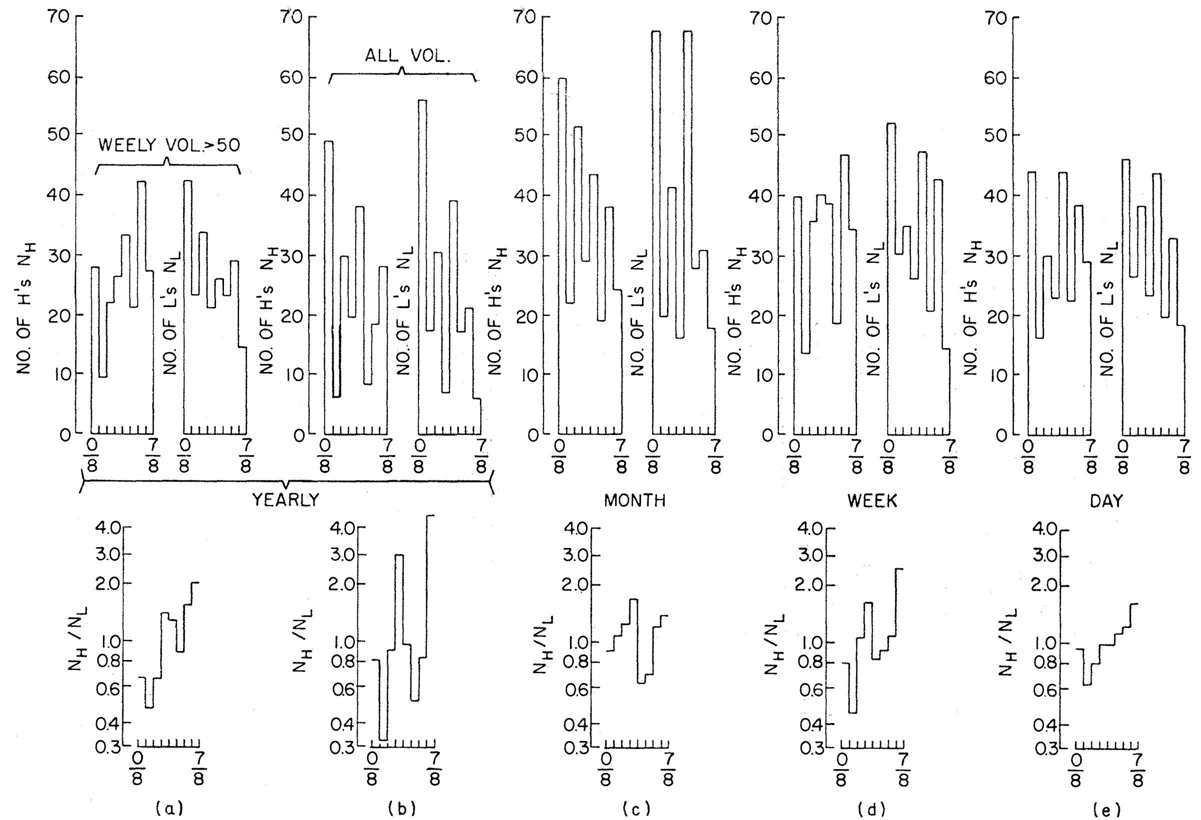

Fig. 12. Distribution of yearly, monthly, weekly and daily highs and lows with respect to eighths, NYSE (common and preferred). The data have been censored; i.e .,the histograms include only stocks for which the close was not equal to either a high or a low for the interval considered. Yearly data 1/1/59 to 1/15/60, data from *Barron's*. (*a*) from E to H, all volume; (*b*) from A to E, weekly volume >50 round lots/week; (*c*) monthly data for June 1960, all volumes, censored fraction =0.20, data from Standard and Poor's *Stock Guide*, NYSE stocks only. Sample from A through C; (*d*) weekly data for week ending 1/15/60, all volumes, NYSE, censored fraction = 0.40. Data from *Barron's*; sample from A through C; (*e*) daily data, 1/27/60, all volumes, NYSE, censored fraction =0.80. Data from *Wall Street Journal*, A through Z.

or 'lows' with true maxima or minima (as defined above), the data of Fig. 12 have been censored, in that we have only plotted in the histogram data for which the close was not equal to a high or low. For such data, one at least of high or low, is in fact a maximum or minimum. This censored-out fraction of the entire sample is quite appreciable for intervals of a day, week, or month. For a year's span of data, it amounts to only a few per cent. The censored-out fraction for a year, as an example, does not correspond to the published list of new highs and lows for the year (exactly an interval of a year only as published at the beginning of the year). To qualify for the censored-out fraction for the year, a stock must close high

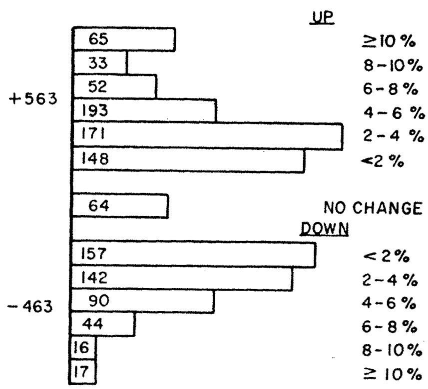

Fig. 13. Histogram of percentage price changes from "The Exchange" for 1090 NYSE listed common stocks for the month ending July 15, 1958.

(low) for the day, and at a price equal to or greater (less) than the published high (low) for the preceding year.

The clustering of maxima and minima is shown in the plots of the ratios N_H/N_L (censored) against the eighth/fraction(Fig. 12). The characteristic behavior of this ratio indicates a tendency for maxima to cluster on the low side of integers and half-integers ($7/8$ and $3/8$, $N_H/N_L > 1$); the minima on the high side of integers and half-integers ($1/8$ and $5/8$, $N_H/N_L < 1$). The clustering effect is more pronounced at the whole numbers than the halves. The effect is also more pronounced the smaller the volume [*compare* Figs. 12(a) and (b)].

If one were to speak in terms of physical analogs, one could say that the diffusion of stock prices took place on a one-dimensional diatomic lattice, in which the preferred sites of occupancy were the integers, and less strongly preferred, the half-integers. The preferred sites also act as 'scattering centers' off which the 'particles' tend to 'bounce.' At increased 'temperature' (cf. volume or dispersion) the importance of these sites as preferred positions and scattering centers tend to diminish.

The preference of human decision-makers for certain numbers over others is not a peculiarity of the stock market. KENDALL[13] gives several examples of this phenomenon, which he calls "one of the most insidious with which psychology has to deal." The phenomenon is slightly different in the stock market, since here the preference is not a consequence of observational bias, or errors. It rather seems as though the collective mind of the market were operating as a binary digital computer, with a built-in preference for round (in the binary sense) fractions.

The existence of the ruler effect is probably responsible for what would ordinarily be regarded as a most unconventional and even amateurish graphical presentation of data. Figure 13, reproduced from "The Exchange" gives a histogram of monthly per cent changes of 1090 common stocks. It will be seen that the members of the sample which fall precisely on the class boundary at zero are treated as a separate class. Conventional practice would have divided these equally between the upper and lower classes, or else moved the class boundaries to odd integral percentages. In this instance the editors of "The Exchange" are to be complimented for faithfully reporting a peculiar phenomenon most unlikely to have occurred by chance (by chance here meaning an equal probability with respect to eighths of closing prices, or no ruler effect). A conventional graphical representation would have obscured the phenomenon.

We can compare the expected (with no ruler effect) and observed number of stocks which fall in the class 'no change' in the following way. Let P_1 be the price of a common stock picked at random at the beginning of the monthly period and P_2 its final price. The distribution of P_1 is approximately given by

$$\Phi(X)\, dX = \exp-[(X-X_0)^2/2\sigma_{X^2}]/(\sqrt{2\pi}\,\sigma_X)\, dX. \qquad (2)$$

Here $X=\log_e P_1$. Using typical figures from reference 1, Figs. 4 and 6, a median priced stock is given by $e^{X_0}=\$35=P_{10}$, and $\sigma_X=0.65$, corresponding to the fact that two-thirds of stock prices fall into the range \$18 to \$66. The distribution of $S=\log_e P_2/P_1$ is given by

$$\Psi(S)\, dS = (1/\sqrt{2\pi}\,\sigma_S)\exp-[(S-S_0)^2/2\sigma_S{}^2]\, dS. \qquad (3)$$

This is essentially what is given by the histograms of "The Exchange" (Fig. 13) with percentages expressed as decimal fractions. σ_S falls in the range 0.04 to 0.07 (Figs. 3 and 4). S_0 is quite closely the deviation, as a decimal, of the median of these histograms; it is usually (as in Fig. 13) less in absolute value than σ_S.

If we make the assumption that X (i.e., the initial price) and S (approximately percentage of change as a decimal, in a month) are independently distributed, then the joint distribution of S and X is

$$\varphi(X, S)\, dX, dS = \frac{e^{-(X-X_0)^2/2\sigma_X{}^2} e^{-(S-S_0)^2/2\sigma_S{}^2}}{2\pi\sigma_S\sigma_X}\, dX\, dS. \tag{4}$$

[Admittedly, this assumption of the independence of S and X, or that "volatility as a per cent is independent of price," is an approximation, since Fig. 2 gives evidence that the dispersion σ_S does depend somewhat on the price class (X) to which σ_S refers. Nevertheless, the results which follow do not depend in any essential way on this approximation. In fact, the integrations both P_1 and P_2 which follow could all be carried out completely numerically, using different dispersions for each price class. Doing this would not alter appreciably the conclusions which we draw.] From the distribution of S in Fig. 13 we obtain the distribution of P_1 and P_2 as follows, using numerical values $\sigma_S = 0.045$, $S_0 = +0.0055$.

$$\phi(P_1, P_2)\, dP_1\, dP_2 = \phi\left(X = \log_e P_1, S = \log_e \frac{P_2}{P_1}\right) \left|\frac{\partial(S,X)}{\partial(P_1 P_2)}\right| dP_1\, dP_2 \tag{5}$$

$$= \frac{e^{-(\log_e P_1 - X_0)^2/2\sigma_X{}^2} e^{-(\log_e P_2 - \log_e P_1 - S_0)^2/2\sigma_S{}^2}}{2\pi\sigma_S\, \sigma_X} \frac{dP_1\, dP_2}{P_1\, P_2}. \tag{6}$$

Then for a sample size $N_I = 1090$ for the number of issues (Fig. 13), the expected number in the class 'no change' of Fig. 13, since prices must move by $\frac{1}{8}$ point, or not at all, is

$$\mathcal{E}[N_{(\text{no change})}] = N_I \int_{P_1=0}^{\infty} \int_{P_2=P_1-1/16}^{P_2=P_1+1/16} \phi(P_1\, P_2)\, dP_1\, dP_2. \tag{7}$$

The integration over P_2 can be executed by taking the value of the integral at the mid-point, P_1, times $\frac{1}{8}$, the range of integration, and the final result is

$$\begin{aligned} \mathcal{E}[N_{(\text{no change})}] &= \frac{N_I}{2\pi\sigma_S\, \sigma_X} \frac{e^{-S_0{}^2/2\sigma_S{}^2}\, e^{+(1/2)\sigma_X{}^2}}{8P_{10}} \\ &= 25, \end{aligned} \tag{8}$$

using numerical values for N_I, S_0, σ_S, σ_X, P_{10} given above. The observed value of 'no change' is 64, which is well outside the expected variation, or dispersion $\pm\sqrt{25}$ of the computed answer. Hence, we conclude that the class 'no change' is much more heavily populated than could be expected by chance; that is, with no ruler effect.

A similar calculation using equation (8) and a figure for the across-the-market dispersion for one 24-hour calendar day can be made for the 'unchanged' class which is published daily in the press. Typical observations indicate that 'unchanged' averages $N_u \cong 200$–250 stocks in 1961, with a sample size $N_I \cong 1200$. The computed figure for N_u with no ruler effect is only about half the observed number.

An analog to the above phenomenon which may appeal to the mechanical-minded, is that stock prices move by a 'slip-stick' mechanism, like motion with friction (resistance) rather than as a 'smooth' response to economic forces. The sticking points are the whole, and to a lesser extent, half integers.

The data of Fig. 12 contain an additional item which allows an interesting comparison with the theoretical prediction given from stock prices considered as simple random walks. The legend of this figure contains

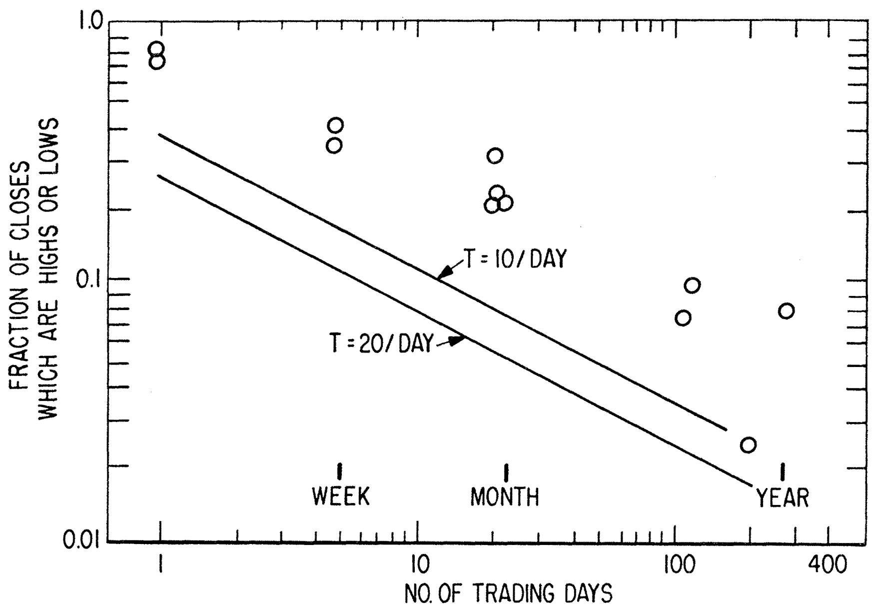

Fig. 14. The fraction of closing prices that are also highs or lows, against the interval of observation. Data from Fig. 12 plus additional data samples from the years 1958–61. The straight lines are the theoretical values for a simple random walk, for 10 or 20 transactions = steps per day.

the censored-out fraction of the total number of stocks considered for which the 'close' price was equal to either the 'high' or 'low' for the interval considered. We have plotted this fraction against the length of the corresponding interval in Fig. 14, and also the theoretical prediction for this fraction under the assumption that there are 10 to 20 steps, or transactions per day. FELLER's formula (reference 12, page 79), in his notation, gives this fraction as $2u_{2n} \cong 2/(\pi n)^{1/2}$, where n is the number of steps. [To derive this result, note his definition of 'first' maximum, equation (8.2), rewrite his equation (4.6) with reversed inequality sign, and apply the reversed path, or 180° rotation argument (p. 70) to both expressions.] It will be seen that the $1/\sqrt{n}$ dependence predicted by theory is fairly well-confirmed by the observations, but the tendency for close prices to be concentrated at highs or lows is even more pronounced than the already very pronounced tendency to do this, which simple random walks exhibit. Why stock prices should have this tendency in such an exaggerated form, we do not know. This subject deserves further examination. We noted in collecting the data for Fig. 14 that the censored-out fraction for intervals of a day, week, month, year contained increasing percentages of inactive and preferred stocks (also of low volume). This emphasizes that the data of Fig. 14 should not be regarded as referring to a homogeneous sample of common stocks, all trading at about the same volume, which the 'theoretical' lines of Fig. 14 tacitly assume.

Since the ruler effect indicating preference for even eighths was so

TABLE IX

OBSERVED NUMBER OF WHOLE NUMBER (NO EIGHTHS) BIDS AND ASKS IN NATIONAL MARKET JUNE 12, 1961

Data from *Wall Street Journal* (first four columns)

Bids	136
Asks	73
Difference	63

Estimated dispersion of difference $= \sqrt{136+73} = 14.3$, under assumption of equal probability of whole numbers in bids and asks. Prob. (Diff. $>63) \ll 0.01$.

pronounced, it was thought that some degree of preference might be manifested in the units place for the whole numbers 0–9. A χ^2 test on a number of NYSE stocks against the hypothesis of equal probability for the ten digits, 0–9, in the units place, gave no significant departures from equal probability, for closes, highs, or lows, for the year, though an even-odd discrimination or a preference for 0 and 5 might have been expected, and might well be found by a more searching examination than we have carried out. However, one interesting side effect was observed. It occurred to us that such a digital effect, if it existed, would be the more pronounced in the over-the-counter market, since here the volume per issue is smallest, and it had already been observed that the ruler effect in eighths increased with decreasing volume. The results of a χ^2 test for digital preference in the units place were ambiguous, but the following curious difference between the bid and ask prices was observed. It amounts to the fact that the ruler effect is more pronounced in the bid than in the ask price. If one examines the frequency of *whole numbers* (no eighths) in the bid and ask prices for one page of over-the-counter quotations from the *Wall Street Journal*, it will be seen (Table IX) that the difference in preference for whole numbers in bid vs. ask is much greater than could be expected on the basis of equal preference. If this effect carries over to the NYSE it means that on the books of the specialists (not normally open to public inspection) the distribution of limit orders to buy below the market is more concentrated at whole numbers, or whole numbers and even eighths, than the corresponding concentration of limit orders to sell above the market. Such a different degree of clustering of orders above and below the market should mean that motion downside is 'jerkier' (more slip and stick) than upside motion. We have not tried to verify this conjecture. It should be noted that the published prices used in Table IX are 'outside' prices, presumably rounded off for public consumption. The 'inside' bid and ask prices between dealers are closer together, and might show the effect less.

It should not be supposed, in presenting the data in this paper, that all the effects we have described are unknown to professionals trading in the market; on the contrary, many of them doubtless are, though we have not previously seen the evidence spelled out, and a significance test applied in cases where the effect was not grossly obvious. For example, Granville (reference 6, p. 162) gives a technique which exploits the barrier effect of whole numbers, and the mechanics of trading and recording data in the 'point and figure' method[14] gives tacit recognition to the 'ruler' and 'barrier' phenomenon.

It would be appropriate to discuss here the relation of the results we have found to previous statistical examinations of stock market and other economic price series. Reference 15 and especially reference 16, give an excellent summary of the difficulties and ambiguities associated with analyzing time series by Fourier and correlogram methods, according to autoregressive schemes, such as were first introduced by Yule. The periodic in time dispersion, and ruler and barrier structure in prices which we have shown, add to the difficulties which Kendall discusses so clearly. These barriers tend to put into the price paths reversals of motion or sequential changes of the sign of price changes. This in itself implies negative serial correlation of price changes, which would tend to subtract from any intrinsic positive correlation associated with the 'inertia' of price motion, or tendency of price 'velocity' to persist, to use terms borrowed from dynamics.

Evidence of negative correlation has been found by MOORE[17] for the first serial correlation coefficient of weekly changes in log prices of individual NYSE common stocks. Kendall[15] found small positive values for the first serial correlation coefficient of weekly changes, but his data referred to British stock indices, so that the two conclusions are not contradictory.

SUMMARY

We would like to summarize our discussion of the internal properties of stock prices by making the following observations. The picture of chaotic, or Brownian motion does not imply that there can be no underlying rational structure. We have tried to show that there is some underlying structure associated with what appears superficially to be the epitome of unrelieved bedlam. There are statistical tests for deciding whether a commonly held belief about the market is in fact justified by the evidence, and there are natural phenomena whose data present problems for analysis quite similar to those offered by the market. The magnitude of the effects we have shown and the way in which they were shown to be significant, do not deny the validity of the premise that the most probable value of the expected change in $\log_e$ price from a *random* choice common stock at a *random* time, is zero.[1] They rather emphasize that under specific conditions and times, one can find a sample of stocks for which $\mathcal{E}(\Delta\log_e P)$ is slightly different from zero. This conclusion is not in conflict with the statement about the most probable value of $\mathcal{E}(\Delta\log_e P)$ used to derive, following Gibbs, the elementary properties of Brownian motion.

The stock market is a gigantic decision-making phenomenon. It deserves scientific attention from those who would like to understand how decision making occurs, naturally, and in the large. As an economic phenomenon, we believe the market can reproduce in a few weeks a scaled version of supply-demand relations that would take many years to complete in a different setting. As a high-speed economic phenomenon it gives unique opportunities for the study of economic behavior.

REFERENCES

1. M. F. M. Osborne, *Opns. Res.* **7,** 145–173 (1959).
2. ———, *Opns. Res.* **7,** 807–810 (1959).
3. W. J. Dixon and F. J. Massey, *Introduction to Statistical Analysis*, McGraw-Hill, New York, 1957.
4. J. C. Clendennin, *J. Finance* **VI,** 398 (1951).
5. C. O. Hardy, *Odd Lot Trading on the NYSE*, Brookings Institute, New York, 1939.
6. J. E. Granville, *A Strategy of Daily Stock Market Timing for Maximum Profit*, Prentice Hall, Inc., Englewood Cliffs, N. J., 1960.
7. H. M. Baker, *Barron's Weekly*, December 5, 1960.
8. A. Merjos, *Barron's Weekly*, p. 5, October 9, 1961.
9. M. F. M. Osborne, *Bull. Amer. Phys. Soc. Series II* **5,** 286 (1960).
10. R. P. Edwards and J. Magee, Jr., *Technical Analysis of Stock Trends*, 3rd ed., Chap. XIII, Stock Trend Service, Springfield, Mass., 1954.
11. F. W. Taussig, *Quart. J. Economics* **3,** 394 (1920–21).
12. W. Feller, *An Introduction to Probability Theory and Its Application*, Vol. 1, 2nd ed., Wiley, New York, 1957.

13. M. G. Kendall, *Advanced Theory of Statistics,* Vol. 1, pp. 187–189, 3rd ed., Charles Griffin, London, 1947.
14. E. J. Staby, *The Chartcraft Method of Point and Figure Trading,* William-Frederick Press, New York, 1949.
15. M. G. Kendall, *J. Royal Stat. Soc. A* **116,** 11 (1953).
16. ———, *Contributions to the Study of Oscillatory Time Series,* Cambridge University Press, 1947.
17. A. F. Moore, Thesis, University of Chicago, Graduate School of Business Administration (1961) (not yet published).