Invariant patterns of wheat price series

Bertrand M. Roehner
LPTHE, Université Paris VII

We show that, despite of their apparent randomness, price fluctuations on different markets follow quite similar and stable patterns. We show that their dispersion remains confined between rather narrow bounds. We analyse the autocorrelation function. It has an exponential decrease and the variation of its relaxation time gives a measure of the economic system's memory. We then show that the distribution of prices follows a lognormal distribution and that the distribution of pairs of prices follows a two dimensional lognormal distribution.

Les séries de prix d'un produit agricole présentent un aspect très chaotique. On montre que l'analyse permet néanmoins de faire apparaître des caractéristiques statistiques permanentes. Ainsi la dispersion du logarithme des prix reste comprise entre des bornes relativement étroites. Le terme dominant de la fonction d'autocorrélation est caractérisé par une constante de temps comprise entre cinq et vingt mois. La distribution des prix et des couples de prix peut être ajustée par une loi log-normale.

I. – INTRODUCTION

1. The scope of this paper

1.1. – Why studying single price series?

In two previous articles (Roehner 1989), which we shall subsequently name article I and II, we were interested in the interdependence between wheat price series on different markets. This interdependence carries, in our opinion, most significant informations about economic exchanges.

In Economies et Sociétés, Série EM n° 11, mars 1990, pp. 5-27

In turning now to analysing single series, our endeavor is mainly one of statistical clarification. This task is however extremely important because it is the first stage required for the construction of any model; indeed, when trying a mathematical description of prices, of their changes and of their interdependence one needs at first knowledge of the characteristics of single price distributions. This is probably the reason why attention was given to the subject by many authors; let us in particular mention : Working (1934), Mandelbrot (1962, 1963), Fama (1965), Granger and Morgenstern (1970), Roll (1970). Fama's article, in particular is a very careful and lucid analysis of the observed distribution of Stock price changes. In section IV, we shall further comment on it.

With the exception of the first, all papers mentionned are concerned with speculative prices namely Stock Market or Commodity Market prices. In this paper we shall consider mainly nineteen-century wheat prices. Speculation already exited at that time : it consisted however in commercial transactions rather than in operations on financial markets. So the speculative component of the prices which will be analysed here is probably smaller than for the prices mentionned above.

1.2. – Which characteristics of price series should be considered?

Instead of simply refering to such standard texts as for instence Burns and Mitchell (1946), Kendall and Stuart (1966), Box and Jenkins (1976), we shall start from scratch again to make the paper almost self contained, mainly for the benefit of economist historians.

First of all, let us again emphasize that we shall consider logarithms of prices as the relevant quantities rather than the prices themselves. This point was explained in article I. In Fama (1965) and in Granger and Morgenstern (1970) the same point of view is adopted

So we shall consider the following time series

$$\bar{p}(t) = \log(t)), t : \text{discrete time variable.}$$

This time-series can be considered as a realization of a stochastic process which we shall denote by $\bar{P}(t)$. Now, what are the relevant characteristics of that stochastic process?

We shall consider three sorts of characteristics :

1) The dispersion characteristics giving an evaluation of the way the process deviates from its average value. See section II.

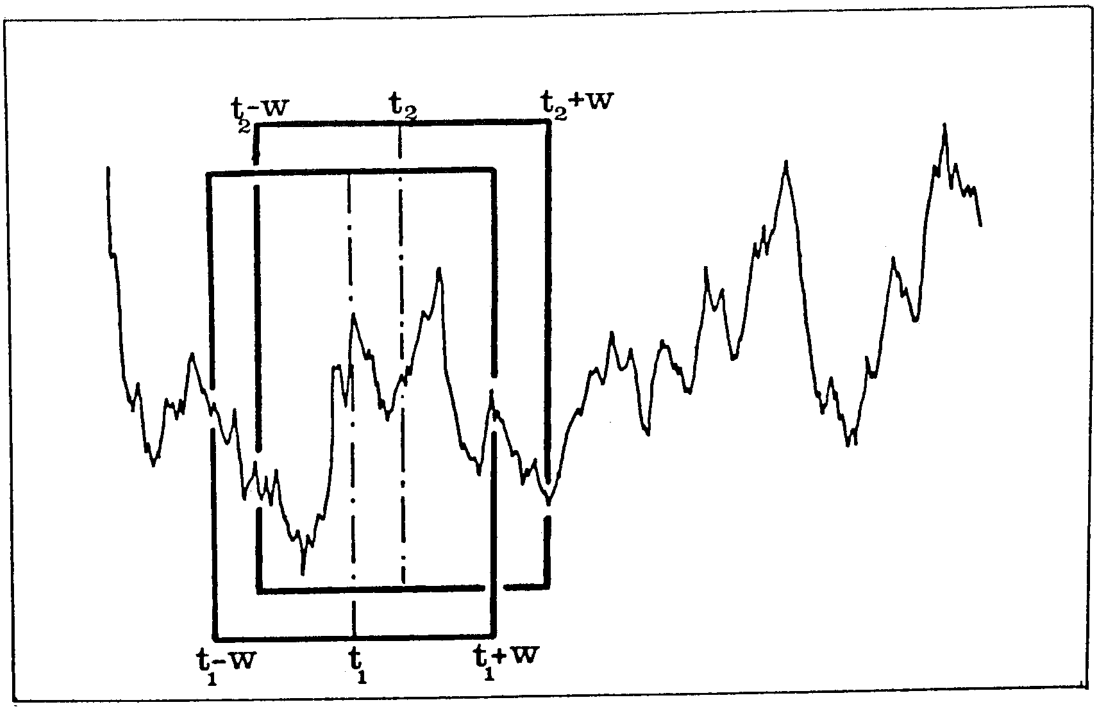
2) The autocorrelation function $A(t)$ which provides an estimation of the memory of the process : the more quickly it goes to zero, the shorter the system memory and vice versa. See section III.

3) The prices distribution function of first and second order along with the distribution of price changes during a given time interval (t_1, t_2). See section IV.

1.3. – The moving window technique

Let us now explain the procedure to be followed to compute those characteristics. We shall use again, as in article I and II, the moving window method; we denote by 2W the amplitude of that moving window. Thus we shall consider, for any time t, the prices falling in the interval $(t - W, t + W)$ (figure 1).

Figure 1
The moving window procedure giving the evolution of numerical characteristics of the time series $p(t)$.



For the variance $D^2(t_1)$ at time t_1, we shall have :

$$D^2(t_1) = \frac{1}{2W} \sum_{t \in (t_1 - W, t_1 + W)} \bar{p}^2(t) - \left[\frac{1}{2W} \sum_{t \in (t_1 - W, t_1 + W)} \bar{p}(t) \right]^2$$

Now, how large should we take the amplitude W ?

For the statistics to be sufficient, we shall require that the moving window contains of the order of hundred prices; to be specific, let us say 300 prices. Let θ denote the periodicity of the time series, that is one day, one week or one month. Thus :

$$2W \sim 300.\theta$$

If θ is one month, we see that W will be of the order of ten years. If we want to observe the evolution of the dispersion over a few decades we see that we need a time series covering at least six or seven decades. We are thus led to discuss the kind of time series that are required here.

2. What kind of time series do we need ?

2.1. – Prices of one product on one market

a) First of all the time series should concern the prices of a *single product on a single place.* Series of price indexes or of price averages over several markets are obtained through various sorts of averaging procedures which could only introduce extra-bias in the analysis of statistical characteristics.

Of course, any observed price is already in a sense time-averaged. Even daily Stock prices represent the average of the different cotations that are performed every day. The kind of averaging that we have in an index is however much more disturbing since many different products are lumped together; thus the short term fluctuations of the index are a complex result of the weighted average of many individual, non synchronized fluctuations. The same can be said for national prices or regional prices that average prices over several markets in the same country or region. The question of the correlation of prices on various markets was discussed at length in article I and II and we saw that even at the beginning of the 20th century, fluctuations on markets distant from more than 100 km present significant differences.

b) The second most important requirement is that we only want to consider market prices. Prices coming from accounting books of private or public institutions usually present slower variations than market prices, because the prices are often the result of medium term agreements between seller and buyer.

2.2. – Stationarity

Another important condition in the perspective of our analysis in stationarity. Indeed we shall determine the characteristics of the stochastic process $P(t)$ through successive time realizations; in other words we assume ergodicity; now stationarity is a necessary condition for ergodicity. This excludes in particular periods of price inflation. Fortunately, the 19th century was a period of stable prices. The inflationary process only started from 1918 on, in the wake of the first World War.

The period 1540-1590 experienced a fast price increase. Since this period is however rather short, the bias it will introduce in the analysis of our wide range series (see table 1) will be limited.

2.3. – Consequences of blancks and requirement as to time coverage

Needless to say, we would like the time-series to have as few-missing figures as possible. It is however possible to somewhat qualify this requirement. Indeed, the computation of the dispersion or of the distribution function is not too much affected by a few gaps. On the contrary, in the calculation of the autocorrelation, even a few blanks would introduce uncontrolable bias. That is the reason why our results for the autocorrelation usually extend over smaller time intervals than the one for dispersion.

As already noted, the time series should cover a period of time at least several times larger than the moving window $2W$. If one considers daily reported Stock Market prices as in Fama (1965), a time series covering a few years can be considered as sufficient. If, on the other hand one considers wheat market prices that are reported every month, one needs time-series extending over at least sixty years.

Table I summarizes the time-series we shall use along with their sources. Besides the series of cereal prices, we mentionned two series giving monthly precipitations. Our objective is to compare them to cereal prices because rain was during the nineteen century and before the main determinant of crop yields (see article II). We shall see that their statistical characteristics are somewhat similar, with the important exception of the autocorrelation : not surprisingly, cereal prices show a much longer memory than precipitation data.

The paper will proceed along following lines :

– in section II, we define the dispersion characteristics we shall use and we explain why they are, in a sense, "universal" characteristics. We shall see that dispersion has a tendency to decrease slowly during the whole 19th century.

– in section III, we introduce the autocorrelation and we show that it can be fairly well fitted through one decreasing exponential. Next, we observe the evolution of the corresponding exponent.

– in section III, we show that the distribution function of the logarithms of prices can be fairly well approximated by a normal distribution. We shall see that this remains true for the distribution of pairs. Hence the distribution of differences of logarithm of prices will be normal too.

To end up, we study the distribution of price changes and compare our results with those of previous authors.

I. – PRICE DISPERSION

1. Dispersion characteristics

1.1. – Definitions

The most obvious dispersion characteristics is the variance and its square-root the standard deviation. The formula for the variance was

TABLEAU I
Price series analysed in this paper

Market	Period	Product	Source
19th Century series			
Bruxelles (Belgium)	1801-1884	wheat	Verlinden (1959) Vol. 3
Bruges (Belgium)	1800-1889	wheat	Verlinden (1959) Vol. 4
Grabow (Germany)	1785-1870	wheat	Beiträge zur Statitik Mecklenburgs (1873)
Parchim (Germany)	1800-1870	wheat	idem
München (Germany)	1790-1855	wheat	Seuffert (1857)
München (Germany)	1790-1855	rye	idem
Wide range series			
Köln (Germany)	1532-1796	wheat	Ebeling, Irsigler (1976)
Köln (Germany)	1532-1796	rye	idem
Köln (Germany)	1532-1796	oats	idem
Köln (Germany)	1532-1796	barley	idem
Paris (France)	1521-1698	wheat	Baulant, Meuvret (1960)
Toulouse (France)	1486-1849	wheat	Frêche (1967)
Precipitation			
Königsberg (Germany)	1848-1923	rain	Clayton (1944)
Utrecht (Netherlands)	1849-1920	rain	idem

already written down in section 1-2, let us write it here in the following alternative form :

$$D^2(t_1) = \frac{1}{2W} \sum_{t \in (t_1 - W, t_2 + W)} [\bar{p}(t) - \bar{m}(t)]^2$$

where $m(t)$ denotes the mean of the logarithms of prices. Another possible characterization is the so-called arithmetical deviation :

$$D_1(t_1) = \frac{1}{2W} \sum_{t \in (t_1 - W, t_2 + W)} |\bar{p}(t) - \bar{m}(t)|$$

It is less popular than the standard deviation because it is not that easy to handle analytically. From a statistical point of view it offers however two advantages :

1) Its evaluation requires less computional work.

2) Its evolution is usually smoother, since it involves the deviations to the power one, whereas the deviations to the power two occur in the standard deviation.

2.2. – Scale invariance

The previous measures have the important property of being independent of the unit of measure. Indeed if one multiplies all prices in the interval $(t_1 - W, t_1 + W)$ by a same number k, the differences $\bar{p}(t) - \bar{m}(t)$ remain unchanged.

Hence, it does make sense to compare the magnitude of $D(t)$ or $D_1(t)$ from one market to another without at all caring about the currency or the cubic measure which was in use; provided of course, that those currency and cubic measure remain the same on each market during the whole period which is considered.

In the same way it makes sense to compare the standard (or the arithmetical) deviation of the logarithms of prices for two periods of the same time-series, where the currency or cubic measure are not the same, provided we leave aside the interval $(t_c - W, t_c + W)$ which contains the transition point t_c between both measurement systems. It is well known by price historians that converting from ancient measurement systems to modern ones is often a troublesome question. The previous remark shows that much can be learned about price series even without that conversion.

2. Statistical observations

2.1. – Magnitude of the standard deviation

Figure 2a depicts the evolution of the standard deviation for the 19th century series. Figure 2b gives the same evolution for the wide range series. The width of the moving window is $2W = 30$ years. It will remain the same subsequently.

a) By mere inspection of both figures, we see that all curves have a very similar evolution and remain confined in a rather narrow interval between 0.20 and 0.40. This is a noteworthy result. This confirms in particular that the rules governing the fixation and registration of market prices were rather similar in Belgium, France and Germany.

b) The dispersion of the logarithms of precipitations falls outside the interval for prices.

Figure 2a
Standard deviation of the logarithms of prices for nineteen century series (curves for Bruges and Parchim, very close respectively to the curves for Bruxelles and grabov, are omitted

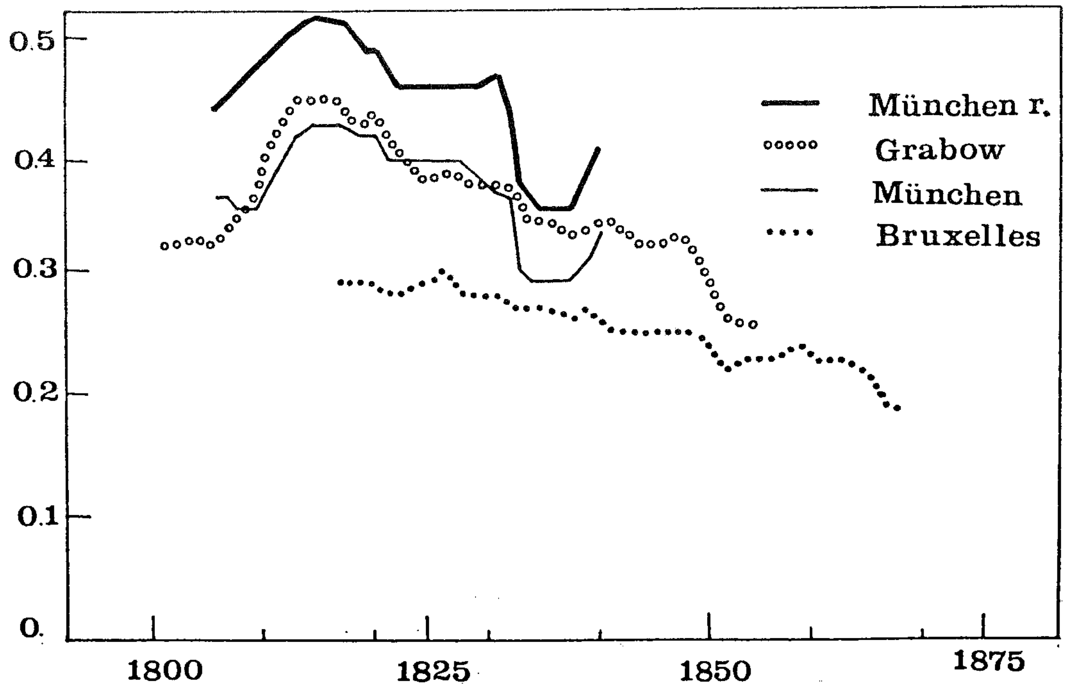
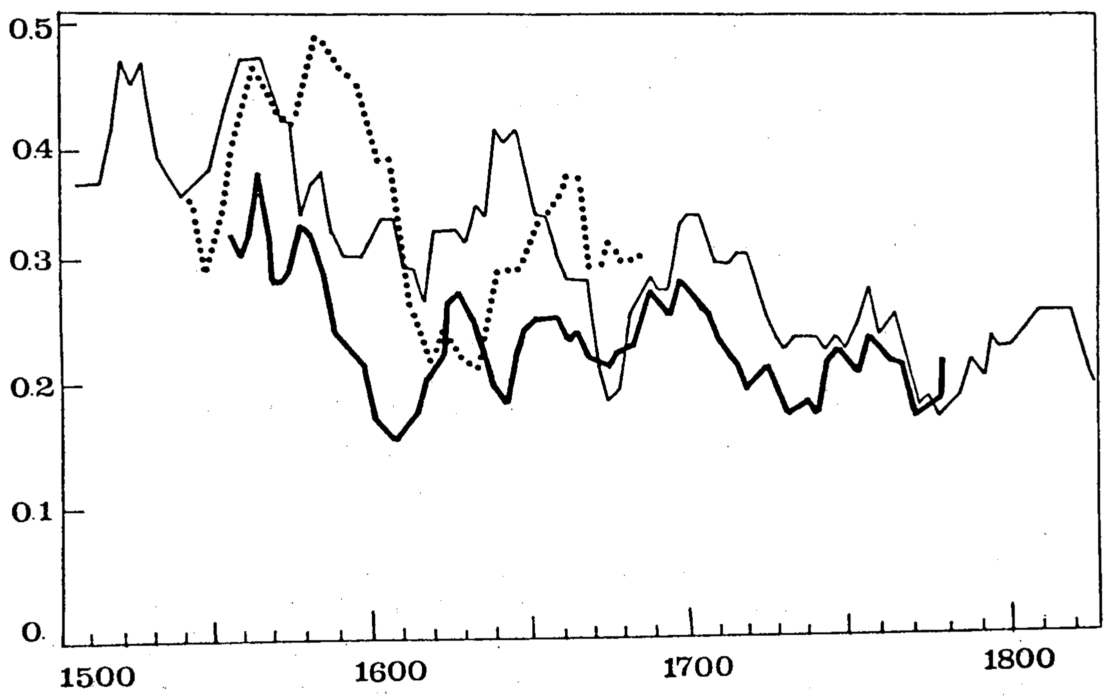


Figure 2b
Standard deviation of the logarithms of prices for wide range series; dots : Paris, thin line : Toulouse, heavy line : Köln



$$0.62 < \text{ s.d. of precipitations } < 0.71$$
in Utrecht
$$0.62 < \text{ s.d. of precipitations } < 0.78$$
in Königsberg

This confirms on a quantitative basis the mere observation that precipitations data are much more irregular from month to month than prices. We shall encounter the same effect in the next section, when computing the autocorrelation. Let us notice that it is again because of the scale invariance mentionned previously, that comparing precipitations dispersion with prices dispersion makes sense.

2.2. – *Dispersion trend*

Let us now consider Figure 2a and 2b separately. Figure 2a shows a slow decreasing trend which is certainly to be attributed to the market integration process which was the main subject of articles I and II. The comparison of the München curve for wheat on one hand and rye on the other, shows that the dispersion was larger for rye. We already explained that point in paper I.

2.3. – *Magnitude of the arithmetic deviation*

Very similar results are obtained for the arithmetic deviation. The variation bounds are the following :

Nineteen century price series (1800 – 1900) $\begin{cases} \text{wheat} : 0.14 < D_1(t) < 0.38 \\ rye \quad : 0.27 < D_1(t) < 0.41 \end{cases}$

precipitations (1850-1920) $\quad 0.47 < D_1(t) < 0.61$

Wide range price series (1500-1820) $\quad 0.11 < D_1(t) < 0.44$

II. – PRICE AUTOCORRELATION

1. Signification of the autocorrelation

1.1. – *Definition*

As already mentioned, the autocorrelation provides a measure of the memory of a stochastic system. Let us see how this results from its definition :

$A(t, t_1)$ = correlation of the

$$\begin{aligned} \text{vector} &: (\bar{p}(t_1 = W), \ldots, \bar{p}(t_1 + W)) \\ \text{and} &: (\bar{p}(t_1 - W + t), \ldots, \bar{p}(t_1 + W + t)) \end{aligned}$$

For a stationary process, the autocorrelation does not depend upon the initial value t_1 ; it depends only upon the time shift t between both vectors.

Let us now consider the evolution of the autocorrelation $A(t)$ with the time shift t : as long as $A(t)$ remains close to 1, it means that the shifted vector $(\bar{p}(t_1 - W + t), \ldots, \bar{p}(t_1 + W + t))$ remains close to the non shifted vector; in other words the process looks very much the same after a moment t. On the other hand, if $A(t)$ drops near to zero, this means that the process changed drastically. The autocorrelation can even become negative meaning then that the two vectors change "a contrario".

1.2. – Exponential decrease

We shall see however that for prices the autocorrelation remains essentially positive, showing a monotonous decrease from 1 toward zero. Thus, a natural idea is to try to fit this curve with a decreasing exponential :

$$A(t) = e^{-t/\tau}$$

The relaxation time τ becomes thus a measure of the memory of the system :
– if τ is small, the system has a short memory
– if τ is large, the system has a memory like an elephant.

2. Statistical results

2.1. – The autocorrelation function

a) Figure 3a shows for instance the autocorrelation function for the München wheat price series in the two years 1806 and 1823. The shape of both curves are clearly compatible with an exponential decrease, except for small times. The values of the corresponding relaxation time are the following, where the correlation coefficient R indicates the goodness of the fit :

1806	$\tau = 8.3 \mp 1.2$ months	$R = 0.92 \mp 0.07$
1823	$\tau = 16.6 \mp 1.0$ months	$R = 0.98 \mp 0.02$

b) Figure 3b represents the autocorrelation curve for precipitations in Königsberg for the years 1867 and 1886. It is clear that these curves

Figure 3a
Autocorrelation function for München wheat prices
heavy line : 1806, thin line : 1823

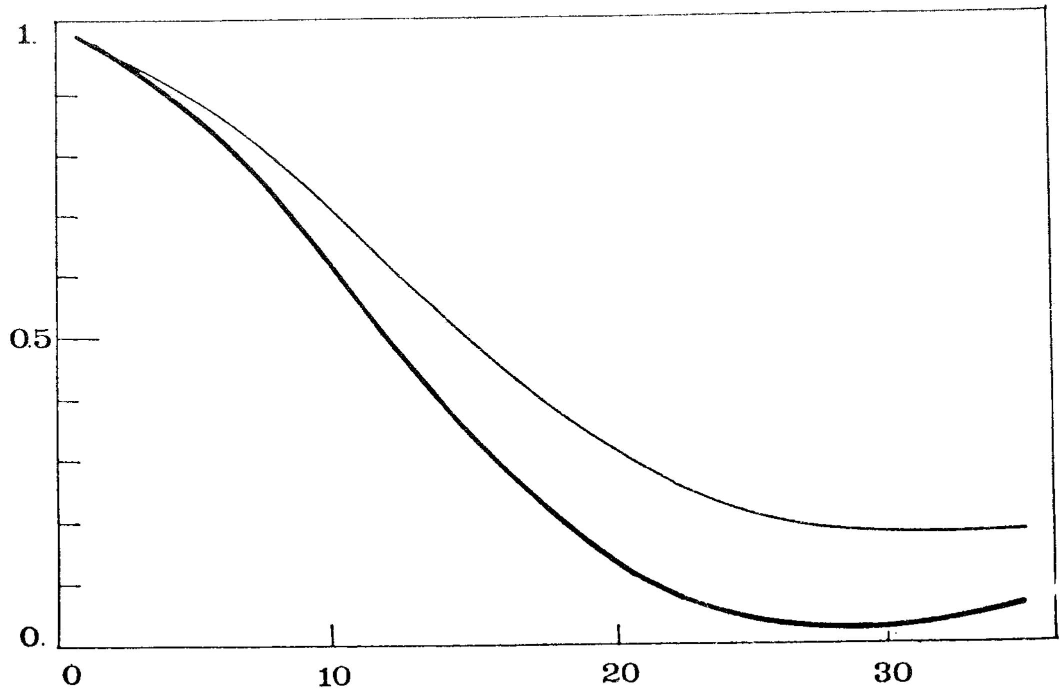
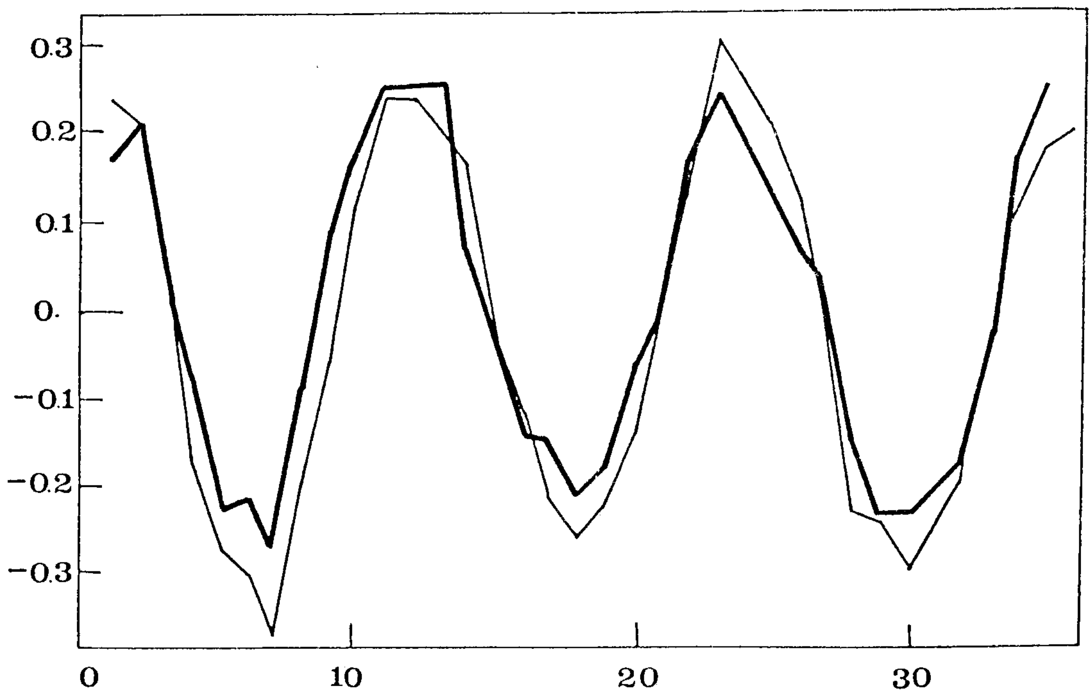


Figure 3b
Autocorrelation function for precipitations in Königsberg
heavy line : 1867, thin line : 1886



are very different from those of figure 3a. The cyclic variation, with period 12 months, has of course its origine in the succession of seasons from one year to another. Now, we are rather interested in the left part of the curve and we observe that its decrease is very fast since the autocorrelation goes from one to zero within 2 months. This confirms our statement that the memory of price series is more than ten times longer than the memory of precipitations series.

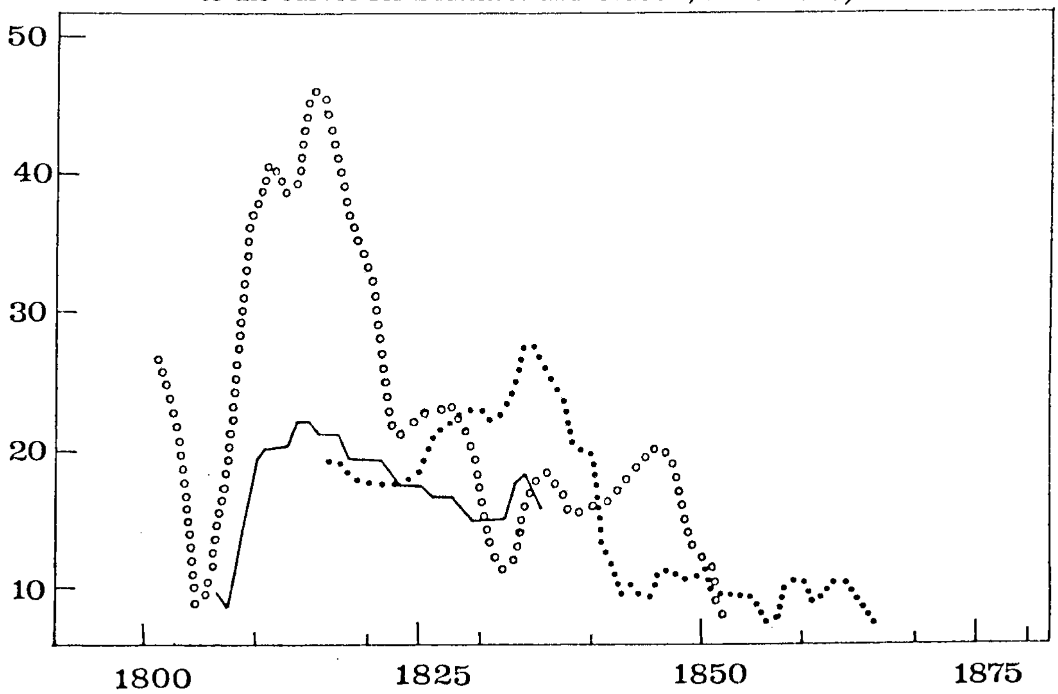
2.2. – Variation of the autocorrelation time

a) Ninteteen century series.

Figure 4 gives the variations of the relaxation time τ for the nineteen century series. The curve for München rye has been omitted since it is very close to München wheat.

Figure 4

Evolution of the relaxation time (in months) for nineteen century series; light dots : Grabow, heavy dots : Bruxelles, thin line : München wheat (curves for Bruges and Parchim, very close respectively to the curves for Bruxelles and Grabow, are omitted)



It is satisfactory that the different curves present parallel evolutions. This conforts us in the idea that these variations are not merely an effect of the price registration procedure. In saying that we have in particular in mind some price series (for instance the series for the market in Eeklo, Belgium given in Verlinden 1959) where some numbers occur repeatedly, this would strongly affect the autocorrelation and we can be sure that such registration bias is absent from the series we are considering.

Now, it is not straightforward to interpret the peaks and lows of Figure 4. The clue to the relation between the relaxation time τ and the economic process remains to be disclosed.

To conclude this paragraph we give in table 2 the mean values of the relaxation time τ and of the correlation coefficient R over the whole period. We gave also the upper and lower bounds of R to indicate its variations range. For τ, we omitted the error bounds : they are comparable to those given in paragraph 1.

b) Wide range series.

We did not represent the variations of the time constant τ for the wide range series; the corresponding curves are still confined to the interval (10 months, 25 months). They display large variations between these bounds without any definite trend.

TABLEAU II

Exponential adjustment of the autocorrelation function

Nineteenth century series	$R(R$ min, R max)	τ(months)
Bruxelles (wheat)	0.93 (0.74, 0.99)	14.5
Bruges (wheat)	0.92 (0.54, 0.98)	10.5
Grabow (wheat)	0.95 (0.79, 0.99)	21.8
Parchim (wheat)	0.95 (0.86, 0.99)	19.9
München (wheat)	0.97 (0.88, 0.99)	16.1
München (rye)	0.98 (0.94, 0.99)	14.6
Wide range series		
Paris (wheat)	0.95 (0.82, 0.99)	17.3
Toulouse (wheat)	0.93 (0.78, 0.99)	11.8

III. – DISTRIBUTION OF PRICES AND DISTRIBUTION OF PRICE-DIFFERENCES

1. Distribution functions

1.1. – Definitions

A stochastic process $P(t)$ is completely characterized by its distribution functions of order n (for any n) namely :

$$f_{P(t_1)\ldots P(t_n)}(p_1, t_1, \ldots; p_n, t_n) = Pr\{p_1 < P(t_1) < p_1 + dp_1 \wedge \ldots \wedge p_n < P(t_n) < p_n + dp_n\}$$

Needless to say, these functions are in general rather complicated; so we shall restrict ourselves to $n = 1$ and $n = 2$, that is to the distribution function of the prices :

$$n = 1 : f_{P(t_1)}(p_1, t_1) = Pr\{p_1 < P(t_1) < p_1 + dp_1\}$$

and to the distribution function of the pairs of prices :

$$n = 2 : f_{P(t_1),P(t_2)}(p_1, t_1; p_2, t_2) = Pr\{p_1 < P(t_1) < p_1 + dp_1 \wedge p_2 < P(t_2) < p_2 + dp_2\}$$

Remark : If the stochastic process is markovian, the distribution functions of order 1 and 2 suffice to completely define the process. It is however almost certain that prices do not form a Markovian process since they have such a long memory as was showed in previous section.

1.2. – Economic implication of the distribution of pairs of prices

From an economic point of view, the pairs distribution is of special interest for the buyer and the seller. Indeed both of them are interested in the possible price changes which may occur in a given time interval. Now the distribution of price differences is merely a consequence of the distribution of the pairs of prices :

$$Pr\{p(t_1) - p(t_2) = d\} = \int_0^{\infty} f_{P(t_1),P(t_2)}(p_2 + d, t_1; p_2, t_2)dp_2$$

2. Logarithms of prices follow a normal distribution

2.1. – Procedure

To test whether a normal distribution can provide a reasonable fit to the distribution of prices, we used the linearization method applied to the cumulative distribution function (Morice and Chartier 1954, p. 105). That is to say, we apply to the cumulative frequencies a transformation (namely $y = Erfc^{-1}(x)$) which would change an ideal normal cumulative function into a straight line. In the new variable, the test for normality is thus changed into a test for linearity for which we use the standard correlation coefficient method.

2.2. – Statistical results

In table 3, we indicate the mean value over the whole period of the correlation coefficient measuring the goodness of fit. It appears to be quite satisfactory. We added in each case the least and the highest correlation coefficient to give an idea of its variation range. Besides, we give the mean value of the adjusted parameters namely the standard deviations σ and the mean $\bar{m}$.

Let us recall that σ occurs in the normal density function in the following form :

$$f_{\bar{p}(t)}(\bar{p}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(\bar{p}-\bar{m})^2}$$

TABLEAU III
Adjustment of logarithms of prices to a normal distribution

	$R(R\min, R\max)$	σ	$\bar{m}$
Nineteenth Century series			
Bruxelles (wheat)	0.989 (0.974, 0.997)	0.25	7.68
Druges (wheat)	0.990 (0.977, 0.998)	0.25	5.81
Grabow (wheat)	0.989 (0.972, 0.997)	0.36	6.99
Parchim (wheat)	0.993 (0.993, 0.998)	0.36	6.88
München (wheat)	0.977 (0.942, 0.995)	0.38	7.38
München (rye)	0.966 (0.930, 0.986)	0.46	6.96
Wide range series			
Köln (wheat)	0.963 (0.943, 0.998)	0.24	5.55
Köln (rye)	0.979 (0.925, 0,997)	0.31	5.24
Köln (oats)	0.989 (0.969, 0.997)	0.27	6.79
Köln (barley)	0.986 (0.953, 0.996)	0.26	7.28
Paris (wheat)	0.972 (0.946, 0.998)	0.36	6.85
Toulouse (wheat)	0.984 (0.936, 0.999)	0.32	6.39

It is not necessary here to study the evolution with time of the standard deviation. This was already done in section II-B.

2. Distribution of prices

As a consequence of the normality of the distribution of the logarithms, the prices themselves will follow a lognormal distribution. This distribution occurs frequently in economics and has been extensively studied (Aitchison and Brown, 1963). Let us recall that its density function reads as :

$$f_{P(t)}(p) = \frac{1}{\sqrt{2\pi}\sigma} \frac{e^{-\frac{1}{2\sigma^2}(\ell n p - m)^2}}{p}$$

3. Pairs of price logarithms follow a two dimensional normal distribution

Now that we know that the distribution of the logarithms of prices is well approximated by a Gaussian distribution, it is natural to wonder if this remains true for the distribution of pairs. If yes, this would be a good indication for the series of the price logarithms to constitute a realization of a Gaussian stochastic process.

3.1. – Procedure

To test the normality of the pairs $(p(t_1), p(t_2))$ were are performing a variable change in the plane $P(t_1), P(t_2))$

$$A = \bar{P}(t_1) \quad \cos\varphi + \bar{P}(t_2) \sin\varphi$$

$$B = -\bar{P}(t_1) \sin\varphi + \bar{P}(t_2) \cos\varphi$$

The angle φ is chosen in such a way that the new random variables A and B become uncorrelated (H. Ventsel 1973, p. 183) :

$$\text{tg } 2\varphi = \frac{2r\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}$$

Now if $(\bar{P}(t_1), \bar{P}(t_2))$ forms a pair of jointly normal variables, A and B will be two independent normal variables, an assertion which will be easily tested using the method of the previous paragraph.

2.2. – Statistical results

In table 4, we indicate the average correlation coefficient R_Z between A and B (which should ideally be zero) along with the average correlation coefficients R_A and R_B characterizing the goodness of a normal fit for variables A and B separately. As in table 2, we give also the lower and upper bounds.

As can be seen, the adjustement is again quite satisfactory. Table 4 is for $t_1 - t_2 = 7$ months; similar results are obtained for other values of $t_1 - t_2$. The last column indicates the correlation coefficient r between $\bar{P}(t_1)$ and $\bar{P}(t_2)$. It occurs in the formula of the two dimensional Gaussian density function in the following way :

$$f_{P(t_1),P(t_2)}(p_1, p_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \exp\left[-\frac{1}{2(1-r)^2} Q(p_1, p_2)\right]$$

$$Q(p_1, p_2) = (p_1 - m_1)^2/\sigma_1^2 - 2r(p_1 - m_1)$$
$$(p_2 - m_2)/\sigma_1\sigma_2 + (p_2 - m_2)^2/\sigma_2^2$$

r is nothing else than the autocorrelation, already considered in section III. Clearly r will approach 1, when the difference $t_1 - t_2$

TABLEAU IV

Adjustment of pairs of price logarithms
to a two dimensional normal distribution for a time interval of seven months

	R_Z	$R_A(R_A\text{min}, R_A\text{max})$	$R_B(R_B\text{min}, R_B\text{max})$	r
Nineteen Century series				
Bruxelles (wheat)	2.10^{-6}	0.992 (0.98, 0.99)	0.991 (0.97, 0.99)	0.75
Bruges (wheat)	10^{-6}	0.993 (0.97, 0.99)	0.991 (0.97, 0.99)	0.65
Grabow (wheat)	10^{-6}	0.988 (0.96, 0.99)	0.988 (0.97, 0.99)	0.75
Parchim (wheat)	10^{-6}	0.991 (0.98, 0.99)	0.992 (0.98, 0.99)	0.71
München (wheat)	2.10^{-6}	0.987 (0.96, 0.99)	0.982 (0.97, 0.99)	0.79
München (rye)	7.10^{-7}	0.976 (0.95, 0.00)	0.979 (0.93, 0.99)	0.81
Wide range series				
Köln (wheat)	10^{-7}	0.984 (0.96, 0.99)	0.988 (0.97, 0.99)	0.75
Köln (rye)	8.10^{-8}	0.982 (0.93, 0.99)	0.988 (0.95, 0.99)	0.73
Köln (oasts)	3.10^{-7}	0.988 (0.97, 0.99)	0.97 (0.96, 0.99)	0.66
Köln (barley)	5.10^{-7}	0.985 (0.94, 0.99)	0.991 (0.94, 0.99)	0.69
Paris (wheat)	3.10^{-6}	0.972 (0.94, 0.99)	0.985 (0.98, 099)	0.73
Toulouse (wheat)	10^{-6}	0.987 (0.93, 0.99)	0.990 (0.97, 0.99)	0.68

goes to zero, hence the smaller $t_1 - t_2$, the more the density function will be peaked.

3.3. – Distribution of pairs of prices

As a consequence of the normality of the distribution of pairs of logarithms, the pairs of the prices themselves will follow a two dimensional Lognormal distribution. Its density function reads :

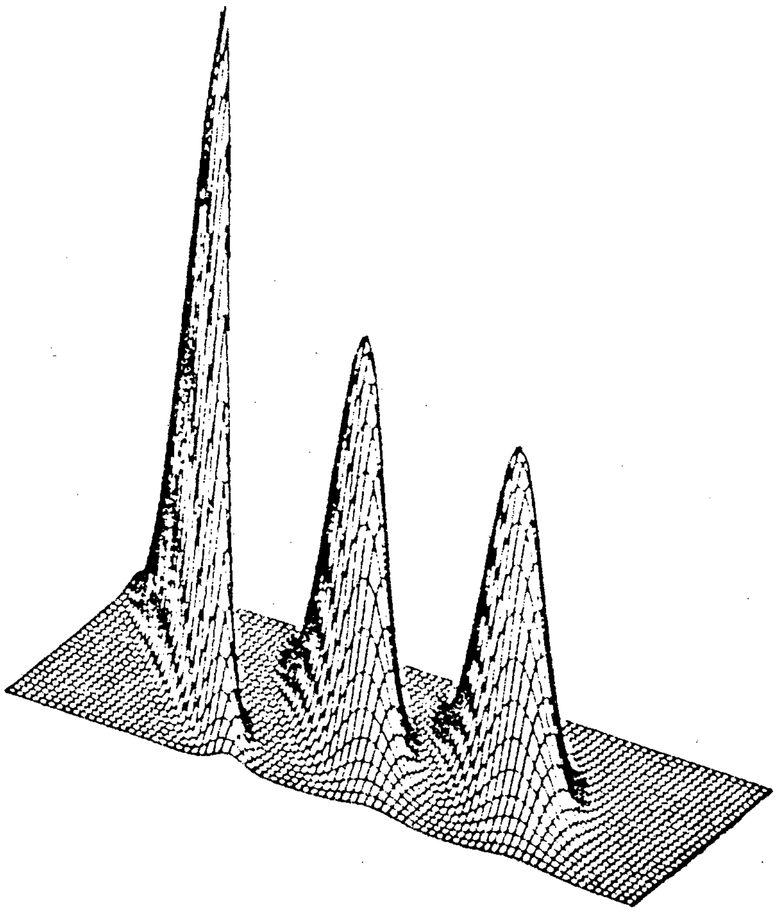
$$f_{P(t_1),P(t_2)}(p_1,p_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \frac{\exp\left[\frac{-1}{2(1-r^2)}Q(\ell n p_1, \ell n p_2)\right]}{p_1 p_2}$$

To give an intuitive perception of the two dimensional distribution, we visualized in figure 5 the surface representing the density function for :

$$t_1 - t_2 = 2 \text{ or } 6 \text{ or } 10 \text{ months}$$

The differences in the shape are quite apparent. The volume limited by each surface is of course equal to 1 : what is gained in height must be lost in breadth.

Figure 5
Representation of the density function of the two dimensional normal distributions of price logarithms (Grabow).
From the highest to the lowest, the surfaces correpond to time intervals of 2 an 6 or 10 months.



4. Distribution of price changes.

4.1. – Distribution of changes of price logarithms

a) From the very beginning, we considered the logarithms rather than the prices themselves as the significant variables; so it is quite natural to also consider the differences of logarithms of prices :

$$\bar{D}(t_1, t_2) = \bar{P}(t_1) - \bar{P}(t_2)$$

If $\bar{P}(t_1), \bar{P}(t_2)$ are jointly distributed normal variables, it is well known that $\bar{D}(t_1, t_2)$ will be a normal variable too with standard deviation given by (Ventsel 1973, p. 267) :

$$\sigma^2(\bar{D}(t_1,t_2)) = \sigma^2(\bar{P}t_1)) - 2r\sigma(\bar{P}(t_1)).\sigma(\bar{P}(t_2)) + \sigma^2(\bar{P}(t_2))$$

Now if $t_1 - t_2$ goes to zero :

1) r will approach 1

2) $\sigma(\bar{P}(t_1)$ will tend to $\sigma(\bar{P}(t_2))$

Thus $\sigma^2(\bar{D}(t_1, t_2))$ will go to zero. This conclusion is indeed confirmed by statistical tests : see figure 7a. It may be summarized in the following rule :

Proposition : The changes on a given time interval (t_1, t_2) of the logarithms of prices are distributed according to a normal distribution; its standart deviation decreases to zero when the interval (t_1, t_2)

Remark : Another way to express the same result is to say that the distribution of the price ratio $P(t_1)/P(t_2)$ is lognormal.

b) Our previous result differs somewhat from the conclusions of Fama 1965. For its Stock price changes, he obtained distributions which are close to normal distributions, with however a slower decrease for both tails than would be expected for a true normal distribution.

The same effect seems to be present here (see figure 5a) but to a smaller extent than for Stock price changes. In other words this deviation only affects a few extreme, and by the way rare, values. Finding the distribution giving the best fit is certainly an important task; but it is as important, or even more, to give that distribution a theoretical basis. Because the normal distribution has so many theoretical advantage and provides on the other hand a pretty good fit, we believe it reasonable first to work in the Gaussian framework and to take into account the previous deviations as "perturbations".

c) Mandelbrot (1963) tried an adjustment of price changes in (log, log) variables which is well suited to Pareto like distributions. To allow comparison with those results we give in figure 5b an example of a (log, log) fit for Brussels wheat price series. The curves of figure 5b are far from being straight lines. That is why a Pareto distribution does not constitute a good candidate for our series, except perhaps for the very tail of the distribution.

4.2. – Distribution of price changes

The buyer and the seller may prefer to know the distribution of the changes of prices themselves. After all, if a wholesale dealer buys ten thousand brushels of wheat at price $p(t_1)$ at time t_1 and sells them at price $p(t_2)$, at time t_2 his profit (or loss) will be given by ten thousands times the difference $p(t_2) - p(t_1)$ and not by the difference of the logarithms!

Figure 5a
Normal adjustment of price changes (Bruxelles) :
cumulative distribution function.
large dots : 1 month interval, small dots : 4 months interval

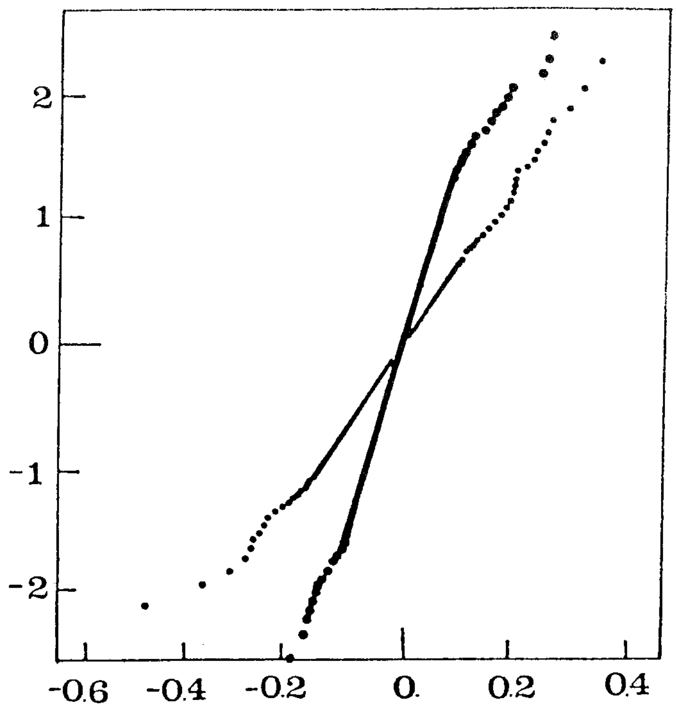
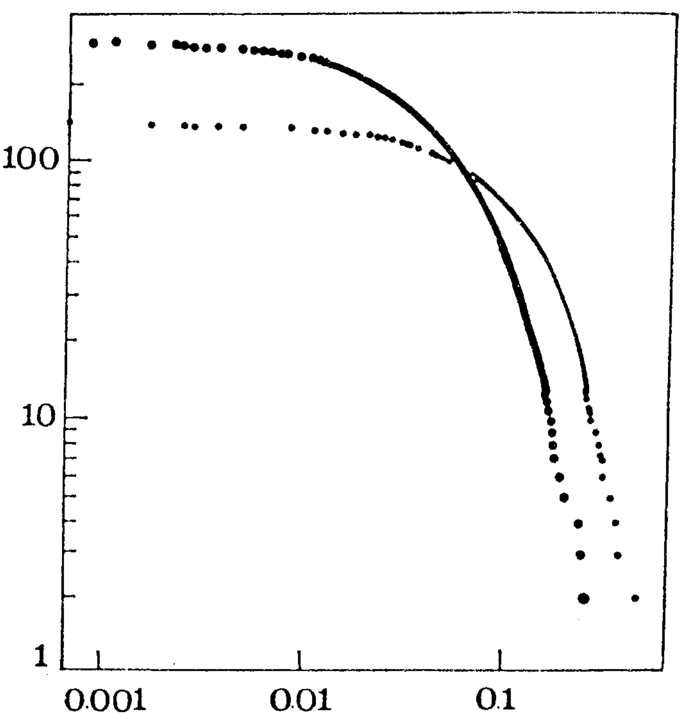


Figure 5b
log-log plot of price changes (Bruxelles) :
number of changes larger than the value indicated on x-axis,
large dots : 1 month interval, small dots : 4 months interval



As already indicated, once the distribution of pairs of prices is known, deriving the distribution of changes is merely a matter of integral calculus. This integral reads

$$f_D(h) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int_0^{\infty} \exp\left[-\frac{1}{2(1-r^2)} Q(\ell n(p_2+h), \ell n p_2)\right] \frac{dp_2}{p_2(p_2+h)}$$

A separate paper will be devoted to it since it is not easily tractable.

CONCLUSION

Our paper was devoted to series of cereal prices. Let us summarize the results we arrived at, concerning dispersion, autocorrelation and distribution of prices :

1) The standard deviation (of the logarithms of prices) remains confined in a narrow interval, namely.

$$0.10 < D(t) < 0.50$$

It shows a slow decreasing trend from 1820 on.

2) The autocorrelation is fairly well approximated by a decreasing exponential of the form :

$$A(t) = e^{-t/\tau}$$

where τ denotes the relaxation time of the autocorrelation function. It varies within following limits :

$$5 \text{ months} < \tau < 25 \text{ months}$$

3) The distribution of the logarithms of prices is fairly well approximated by a normal distribution. This is true as well for the distribution of pairs of price logarithms. A Gaussian stochastic process thus appears as a reasonable candidate for the modelling of wheat price series.

From an empirical point of view several questions remain open; for instance :

Do we have similar results for other goods : other agricultural products or even industrial prices ?

In which way are these results altered in the case of more speculative prices such as for instance 20th century Commodity Market prices ?

Now, as already mentioned in the introduction, the empirical outline given in this paper should be considered as being the first step in the construction of a model where interacting markets will be represented by interdependent Gaussian stochastic processes.

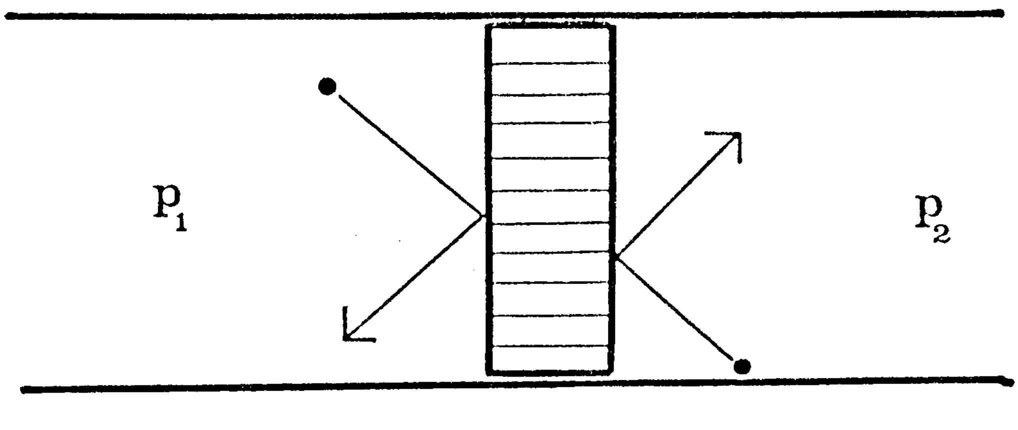
APPENDIX
THE MOVING PISTON MODEL

One could be surprised that we did not investigate the autocorrelation function for price changes. In the case of Stock Market prices this is considered as an important test (Fama, 1965) of the random walk model which considers successive transitions as independent; the validity of that model in turn implies that past prices cannot be used to predict future prices.

When computed from our data, the autocorrelation of price changes between successive months indeed reveals a very fast decrease to zero which seems to support the random walk model as in the case of Stock Market prices. We do not present this result in more details, because we think that it is not of great significance. To explain this point let us present what we shall call the "moving piston model".

We consider a piston moving inside a cylinder (figure 7). The pressure of the gas on the left (respectively on the right) of the piston is p_1 (respectively p_2). We assume that $p_1 > p_2$. The pressures p_1 and p_2 are supposed to be independent of the position of the piston. The movement of the piston is thus a translation from left to right with constant acceleration.

Figure 7
The moving piston



Now suppose it could be possible to observe the movement of the piston between successive molecular impacts on each side of the piston. The fact that $p_1 > p_2$ will remet in the fact that during a very short time interval (t_1, t_2) (of the order of 10^{-8} seconds), the probability for the piston to experience a shock from the left will be larger than for a shock from the right. In other words, the movement of the piston will be a succession of jumps, alternatively on the right and on the left but on average more often to the right than to the left.

We performed a numerical simulation of this experience, and we computed the autocorrelation function of successive position changes. Now this function has a fast decrease to zero exactly as the price autocorrelation function. Nevertheless the movement of the piston is perfectly predictable *on a sufficiently large time scale.*

What we would like to point out is not of course that Stock prices or wheat prices are as predictable as the piston's movement, but rather that the autocorrelation of successive changes does not constitute a convincing test for the random walk hypothesis; a better test would be to consider the autocorrelation of the prices themselves (which was given in section III) or the autocorrelation of changes on a sufficiently long time interval. We shall come back to that question in greater detail in a separate publication.

REFERENCES

AITCHISON J., BROWN J.A.C., *The lognormal distribution with special reference to its uses in economics*, Cambridge at the University Press, 1903.

BAULANT, M., MEUVRET J., *Prix des céréales*, extraits de la Mercuriale de Paris (1520-1698), S.E.V.P.E.N. Paris, 1960.

BEITRÄGE ZUR STATISTIK MECKLENBURGS 1873, Vol. 7, part 3 and 4.

BOX G.E., JENKINS G.M., *Time Series analysis, forecosting and control Holden Day*, San Francisco, 1976.

BURNS A.F., MITCHELL W.C., *Measuring business cycles*, N.B.E.R., New York, 1946.

CLAYTON H.H., *World weather records*, Smithsonian Institution, Washington, 1944.

EBELING D., IRSIGLER F., *Getreideumsatz, Getreide und Brotpreise in Köln*, Mitteilungen aus dem Stadtarchiv von Köln, Heft 65, Köln, 1976.

FAMA F.F., «The behaviour of Stock-Market prices», *J. of Business 38*, 1965, pp. 34-105.

FRÊCHE G., FRÊCHE G., *Le prix des grains, des vins et des légumes à Toulouse (1486-1868), Presses Universitaires de France, Paris, 1967.*

GRANGER C.W.I., MORGENSTERN O., *Predictability of Stock Market prices*, Heath Lexington Books, 1970.

KENDALL M.G., STUART A., *The advanced theory of statistics*, Vols I, II, III, Griffin, London, 1966.

MANDELBROT B., *Sur quelques prix spéculatifs : faits empiriques et modèle basé sur les processus stables additifs de Paul Lévy*, Comptes rendus de l'Académie des Sciences, 254, 1962, pp. 3968-3970.

MANDELBROT B., «The variation of certain speculative prices», Journal of Business 36, 1963, pp. 394-419.

MORICE E., CHARTIER F., *Analyse Statistique, Paris, 1954.*

ROEHNER B., «An empirical study of price correlation», *Env. and Planning A, 21* 1989, pp. 161-173 and 289-298.

ROLL R., *Behaviour of interest rates : the application of the efficient market model to U.S. treasury bills*, New York, 1970.

SEUFFERT G.K.L., *Statistik des Getreide und Viktualien-Handels im Königreiche Bayerm*, Wisz, München, 1857.

VENTSEL H., *Théorie des probabilités*, Mir, Moscou, 1973.

WORKING H., «A random-difference series for use in the analysis of time-series», *J. of the American Statistical Association 29*, 1934 pp. 11-24.