INSTITUTE OF MATHEMATICAL ECONOMICS

WORKING PAPERS

No. 280

Stock Price Clustering and Numerical Perception

by

Andreas Uphaus, Bodo Vogt, Wulf Albers

January 1998

University of Bielefeld

33501 Bielefeld, Germany

Abstract

The stock price clustering phenomenon has been studied in the context of numerical perception and response as described by the theory of prominence. The predictions of the numerical response process which is a model of how persons generate numerical responses if they have diffuse numerical information (know a range of reasonable alternatives) are compared with empirical data. The data consisted of the inside quotes of 30 stocks traded at the IBIS computer exchange in March/April 1993. The principle structure of the data is predicted by the numerical response process of the theory of prominence. A similarity between the quotes and a laboratory experiment in which the diffuse numerical information has been controlled is observed.

Introduction

The phenomenon of price clustering has first been observed at the NYSE (New York Stock Exchange) by M.F.M. Osborne (1962). Price clustering describes the fact that certain multiples of the tick size (the smallest permitted price unit at stock markets or perhaps the smallest currency unit) are more frequent than others. For example at American stock markets 8/8 is more frequent than 4/8, 4/8 is more frequent than 2/8 and 6/8, and 2/8 and 6/8 are more frequent than 1/8, 3/8, 5/8 and 7/8. This is also denoted as: the even eighths (2/8, 4/8, 6/8 and 8/8) are more frequent than the odd eighths (1/8, 3/8, 7/8 and 7/8). M.F.M.Osborne (1962) observed this form of clustering and that clustering increases with decreasing transaction frequency on stock markets. The same phenomenon was observed by V. Niederhoffer (1965). The effect is not limited to stock markets. It was also observed in the Corn and Soybeans Futures Markets (R. Stevenson, R. Bear 1970) and in the Gold Market in London (C.A. Ball, W.N. Torous, A.E.Tschoegl 1985). L. Harris (1991) showed that clustering occurs within different periods of time, for different market structures and for different stocks.

An intense discussion about price clustering started in 1994 when W.G. Christie, J.H. Harris and P.H. Schultz (1994) interpreted the fact that even eighths are more frequent than odd eighths at the NASDAQ (National Association of Securities Dealers Automated Quotation System). They stated that this is based on a tacit collusion between traders to raise profits.

Further examination of the clustering gave more detailed insights in its dependence on the trading rules (P.E. Grodek 1996), market structure, stock characteristics, volatility and trading volume (S.J. Grossman, M.H. Miller, K.R. Cone, D.R. Fischer, D.J. Ross 1997). Besides technical effects the importance of psychological effects has been realized (V. Niederhoffer 1965).

In this paper the relation between the stock price clustering and the perception of numerical information is discussed. The process how persons create a numerical response if they have diffuse numerical information (know a range of reasonable alternatives) is described by the theory of prominence (W. Albers 1997). Due to this process some numbers are more likely to be given as response than others. Assuming that people trading at stock markets do not have exact information, but diffuse information, about the price of a stock the process of generating a price response should be similar to the one proposed by the theory of prominence. Therefore first a short introduction into the theory of prominence is given which is focused on the process how a numerical response is created if persons have diffuse numerical information. Then a laboratory experiment testing this process is described and the results are given. In the main part the predictions of the theory of prominence and the results of an laboratory experiment are compared with the quotes at the German IBIS stock market from March/April 1993. The data contain all quotes and not only the prices. An indicator characterizing the diffuse numerical information proposed by the theory of prominence is the exactness of a number (a set of numbers). The relation between the exactness and the spread is discussed .

The theory of prominence

In the theory of prominence (G. Albers, W. Albers 1983; W. Albers 1997) it is stated that some numbers are easier accessible to persons than others. The easiest accessible numbers are the prominent numbers P which are also called the full step numbers. These are the powers of 10, i.e. 10^z ($z \in \mathbf{Z}$), their halves, i.e. $5*10^z$ ($z \in \mathbf{Z}$), and their doubles, i.e. $2*10^z$ ($z \in \mathbf{Z}$). The **set of prominent numbers** is:

$$P = \left\{ n \cdot 10^z \middle| n \in \{1,2,5\}; z \in Z \right\} = \{\ldots,0.1,\ 0.2,\ 0.5,\ 1,\ 2,\ 5,\ 10,\ 20,\ 50,\ 100,\ 200,\ldots\}.$$

All other numbers x in the decimal system are presented by means of these prominent numbers. The **presentation of a number** x is: $x = \sum_{p \in P} a_p \cdot p$ with $a_p \in \{-1,0,1\}$. For example 17 is presented as 20-5+2.

The exactness of a presentation of a number is the smallest prominent number with coefficient $a_p \neq 0$. In the example the exactness of the presentation is 2. The smaller the prominent number with coefficient $a_p \neq 0$ is, the finer is the exactness. The bigger the prominent number with coefficient $a_p \neq 0$ is, the cruder is the exactness.

Because the presentation of a number is not unique **the exactness of a number** is defined as the crudest exactness of all presentations of a number.

A number has a **level of exactness** if its exactness is equal to or cruder than the level.

The next definition characterizes the exactness of a set of data:

The **exactness of a set of data** Γ is the crudest exactness for which it holds: at least 75% of the data have the exactness Γ or a cruder one.

Numerical response process (NRP) for the decimal system

In this part the situation is described in which persons are asked to give a numerical response, for example if persons answer the questions: "How many inhabitants has Cairo?" or "What is the value of a stock?" In both cases persons know a range of reasonable alternatives (the range between the lowest and highest reasonable alternative). This range might be of different size. The question is which number do they select. Knowing a range of reasonable responses, but not an exact number is denoted as having diffuse numerical information (for more details see W. Albers 1997). The mental process that creates a numerical response given a stimulus (for example diffuse numerical information) is characterized in (W. Albers 1997) as:

"1. A person cannot judge whether the distance between response and stimulus is high or low.
2. The person can judge which of two responses x and y is nearer to the stimulus than the other.
3. The judgement permits the four responses "response x", "response y", "there are equally strong arguments for x and y" and "cannot say" (because the limits of the person's ability of judgement are reached.
4. The judgement is knife-edged, i.e. the judgement "equally strong arguments" nearly never occurs."

The process how persons create a numerical response if they have diffuse numerical information is given by the following process for the decimal system (W. Albers 1997):

start: select a sufficiently large prominent number p, set x=0

step: decide whether the signal is nearer to x or to x+p

  if nearer to x then p=+[p/2][1] , repeat step

  if nearer to x+p then x=x+p, p=[-p/2], repeat step

end: if x, x+p equally strong then respond x+p/2

  otherwise respond x, or x+p

In the loop of the process (called step) the interval in which the reasonable answer lies is halved by finding out to which boundary the answer is nearer. In the process the loop is iterated until the exactness reaches the boundary of the ability of the person to judge. This determines the level of exactness on which the answer is given. Depending on the ability to judge different levels of exactness are reached. Depending on the level

[1] the brackets denote that the exact value of p is not p/2, but the nearest prominent number to p/2 because halving a prominent number does not always lead to a prominent number. This also holds for p=-p/2.

of exactness some numbers are possible as answers and others are not. This will lead to different frequencies of the numbers as answers to questions concerning diffuse numerical information. In figure 1a this is shown for the interval $[0,10]$[2] with all possible halvings up to the level of exactness of 1. In the figure a $^+$ or a $^-$ is added to the numbers (for example 2^+ and 2^-). A $^+$ indicates that this number is the lower boundary of the interval obtained by iterated halving and a $^-$ indicates that it is the upper boundary. All numbers on a higher level of exactness are more easily accessible.

In figure 1b the same process is given for the eighths. In the octal system the prominent numbers between 1 and 8 are: 1, 2, 4, 8. In general these are the powers of 8, their doubles and their halves. In the octal system the process is easier than in the decimal system, because halving of a prominent number always leads to another prominent number. Roundings do not occur. In the figure the resulting numbers are divided by 8. This gives the NRP for the perception of eighths which are used on American stock exchanges.

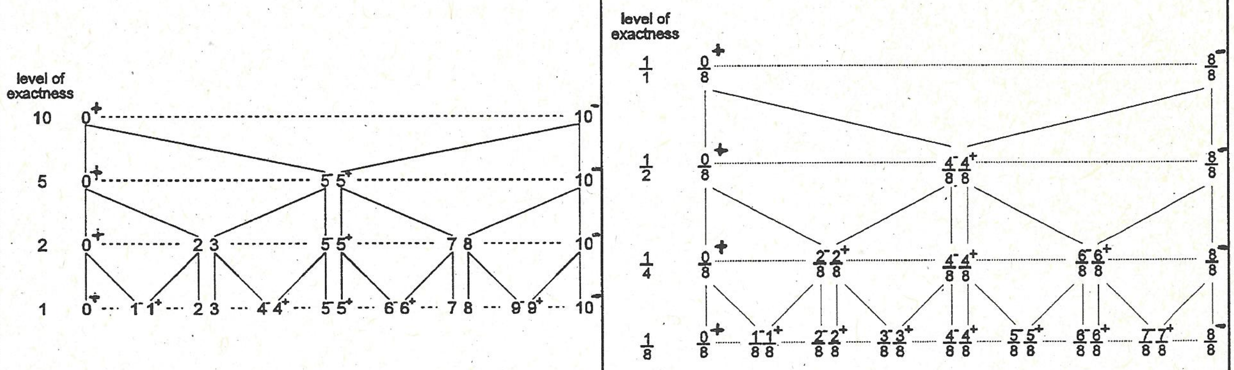
As it can be seen from the figure all even eighths are easier accessible than the odd eighths and 4/8 occurs before 2/8 and 6/8.

[2] These numbers are relevant for the analysis of the German stock prices traded at the IBIS, because the last digits of the stock prices are the multiples of the smallest unit which was 0.1 DM at that time. These multiples are between 0 and 10.

Figure 1: The numerical response process

Figure 1a: **The numerical response process in the decimal system**

Figure 1b: **The numerical response process in the octal system**



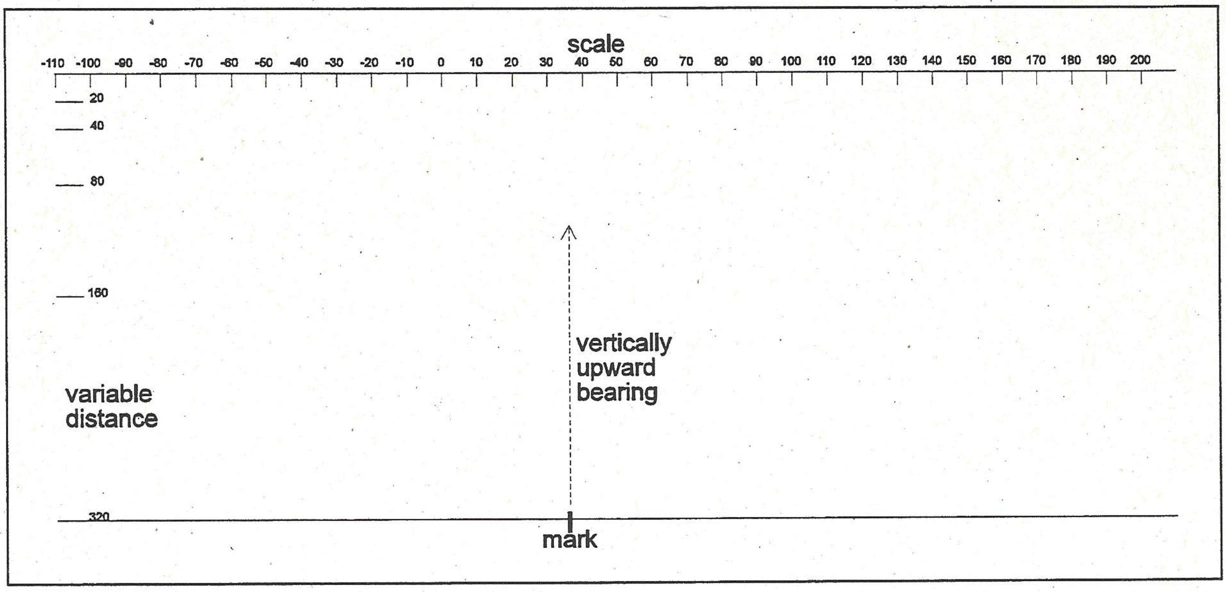
Interpreting the process of pricing a stock as giving a response for diffuse numerical information it is obvious why even eighths are more frequent than odd eighths: because they are easier accessible.

A bearing experiment

The numerical response process has been tested in a laboratory experiment (B. Vogt, W. Albers 1992). In the experiment the subjects saw a computer screen as shown in figure 2. On top of the screen was a scale from -110 to 200 on which all multiples of 10 were marked. The corresponding values (-110, -100,...,190, 200) were displayed above the marks. In a given distance from the scale was another line with only one mark. The distance between scale and lower line was varied (0, 20, 40, 80, 160 and 320 pixels as indicated in the figure). 22 subjects had the task to identify the position of the mark on the scale by bearing vertically upward and answering the point on the scale that was obtained. They were asked to give the response with which they were most content. The position of the mark was varied such that the last digits of the correct positions (0, 1, 2,..., 7, 8, 9) had the same frequency for every distance. The

first two digits of the position of the mark (for example 15 in 154) were selected in a way that the difference between two subsequent positions was essential.
The question examined in the experiment was: How do the frequencies of the last digits depend on the distance between mark and scale? A priori every last digit had the same chance to be given as a response.

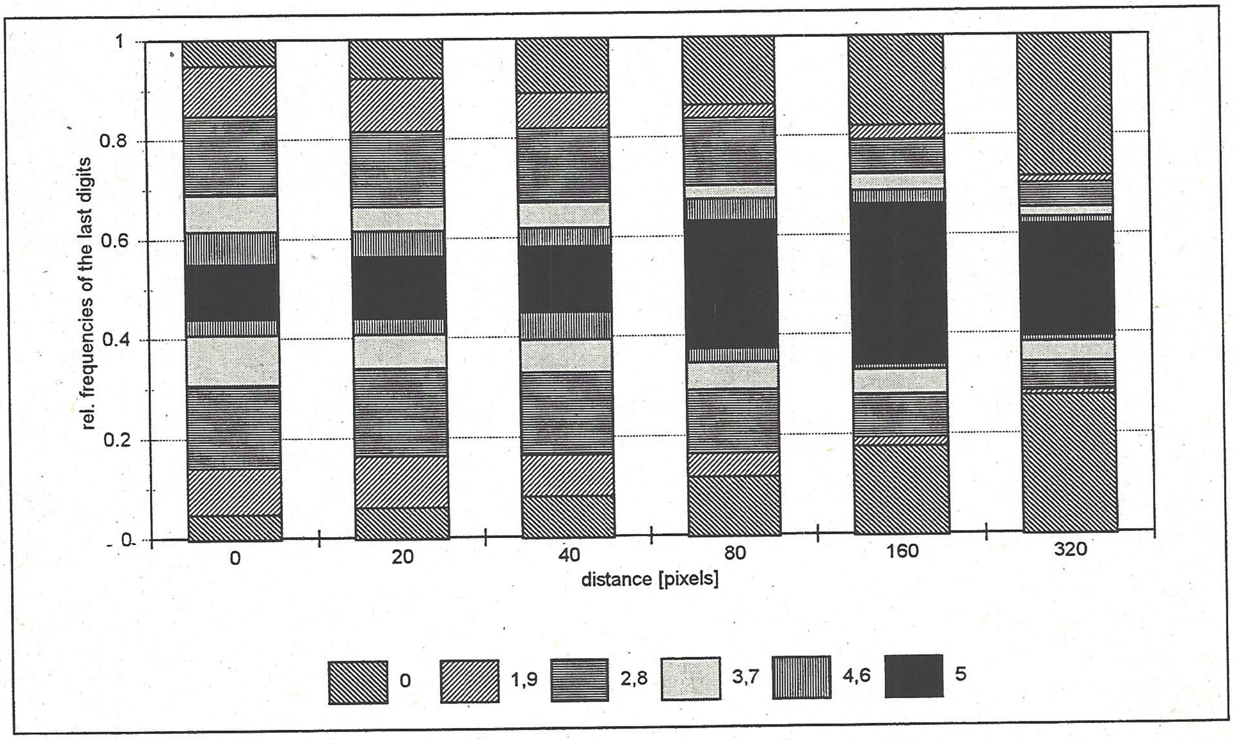
Figure 2: Computer screen used in the bearing experiment



The prediction of the theory of prominence is the following. The bearing process creates an unprecise image of the given signal which is more unprecise if the distance between mark and scale is large. The exactness of the answer should become finer if the distance is reduced. By geometrical considerations it can be obtained that the range of reasonable alternatives is proportional to the distance between mark and scale. Varying the distance between mark and scale gives control over the diffuse numerical information (the range of reasonable alternatives). According to the numerical response process shown in figure 1 the prediction is that for large distances between mark and scale the last digit should be 0. By decreasing the distance first 5 should occur as additional number. Further decrease of the distance should result in responses of 2,8. 2,8 should occur at larger distances than 3,7, because all multiples of 10 are marked on the scale and 2,8 are constructed via 10 (8=10-2 and 2=0+2), whereas 3,7 are constructed via the non marked 5 (7=5+2 and 3=5-2). The same argument holds for an appearance of 1,9 at larger distances than 4,6: 1,9 are constructed via 10 (9=10-1 and 1=0+1), whereas 4 and 6 are constructed via 5 (6=5+1 and 4=5-1). After 2,8 a decrease of the distance between mark and scale leads to responses of 3,7 and 1,9. Finally 4,6 should be given as responses.

The result of the experiment is shown in figure 3. The relative frequencies of the last digits are plotted versus the distance between mark and scale. The numbers can be identified by their hatches. The relative frequencies of the different numbers have maxima at different distances. These sequence of the maxima indicates the sequence of appearance. The appearance of the numbers is as predicted by the numerical response process (NRP) of the theory of prominence.

Figure 3: Results of the bearing experiment



Financial markets

Stock markets and diffuse numerical information

The market price of a stock is influenced by several factors. Among them seem to be uncertainty about the future and incomplete information. Uncertainty about the future is given, because investors do not know for sure how a firm and the corresponding stock will develop. Incomplete information seems to be given, because not all information that is needed for pricing a stock exactly is available to all investors. For example all the last activities of a firm are not exactly known, neither all of the newest technologies, neither the exact economic fundamentals. All these facts should lead to a diffuse numerical information about the price of a stock. Also technical analysis will often lead to a range of prices - at least if different approaches are applied. The numerical signals persons get at the stock exchanges are the last prices, bids and asks which vary during a day indicating that the price is not known exactly. Therefore in principle the numerical response process for diffuse numerical information should influence the stock price. The situation seems to be similar to the one described in the part about the numerical response process: persons can judge which price is nearer to the signal, but only up to a certain limit given by the boundary of their ability of judgement determined by the uncertainty and the incomplete information.

Other factors that influence the price are strategic considerations in the context of order placements. These have been examined in a different study (W. Albers, A. Uphaus, B. Vogt 1998).

The data of the financial markets

The data consist of 194068 quotes placed inside the spread of 30 DAX stocks (the names of the stocks are shown on the x-axis of figure 6) traded at the IBIS (in Frankfurt) in March/April 1993. The quotes that are placed inside the spread are denoted as inside quotes.

In contrast to laboratory experiments the diffuse numerical information cannot really be controlled. This problem is tried to solve by considering the exactness of the stock prices which gets finer if the ability of persons to judge increases. An exactness of a stock is determined by taking all prices in the period and applying the rule for the exactness of a set of data (the exactness of a set of data Γ is the crudest exactness for which it holds: at least 75% of the data have the exactness Γ or a cruder one) to this set. With this exactness the spread (next part) and the last digits[3] of the stock prices are compared.

Exactness and spread

The size of the market spread can be regarded as an indicator how exact the stock price is known. For the exactness defined in the theory of prominence it holds: the finer the exactness of the stock price the further did the numerical response process proceed. Due to these arguments the expected relation is that the more exact the prices are the smaller the spread should be.

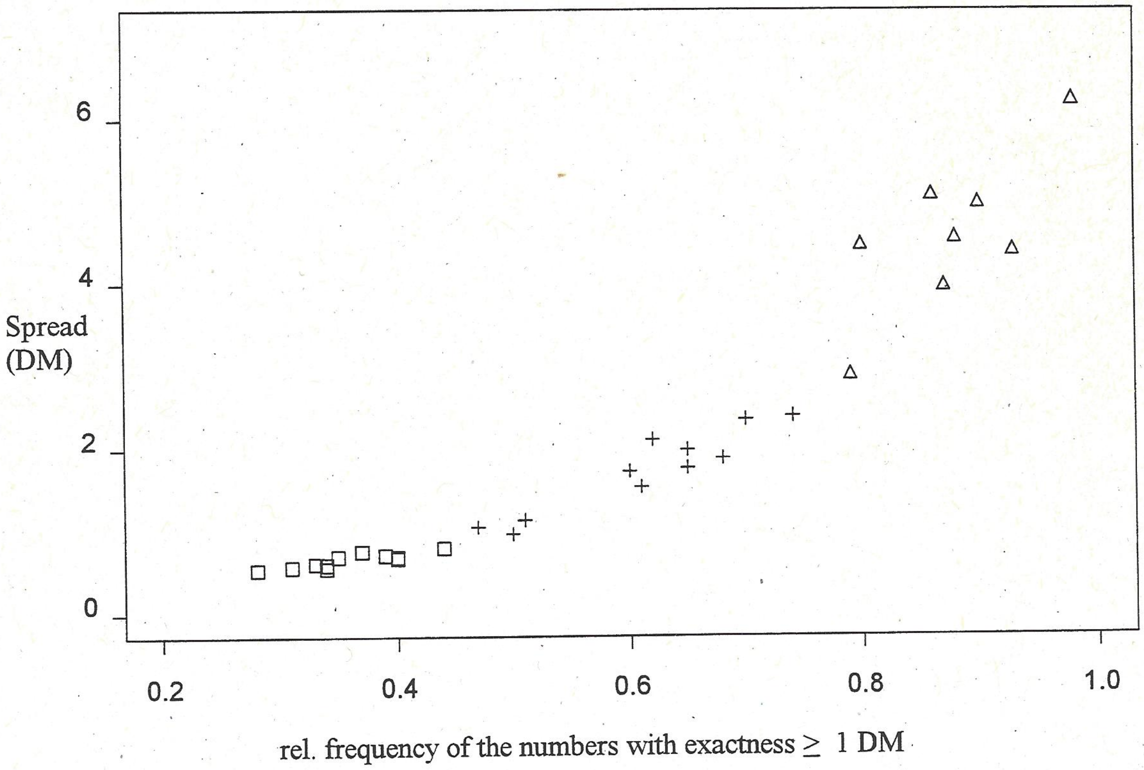
Looking at the relation between the exactness Γ of prices and the spread leads to the result shown in figure 4. The symbol of the stocks with exactness 0.2 DM is a □, the symbol for the stocks with exactness 0.5 DM is a + and the symbol for the stocks with exactness 1 DM is a Δ . Because these data are on the aggregate level a further differentiation of the stocks is obtained by ordering them according to the frequency of the prices with the crudest exactness (of 1 DM[4]). This ordering reflects that sometimes prices are on a cruder or finer level of exactness than the determined level of exactness due to the perception of some traders. It is shown in the figure that the spread increases as the exactness of the stock prices get cruder.

---

[3] For the German stocks the last digit is a multiple of the smallest unit (of 0.1 DM)

[4] or cruder

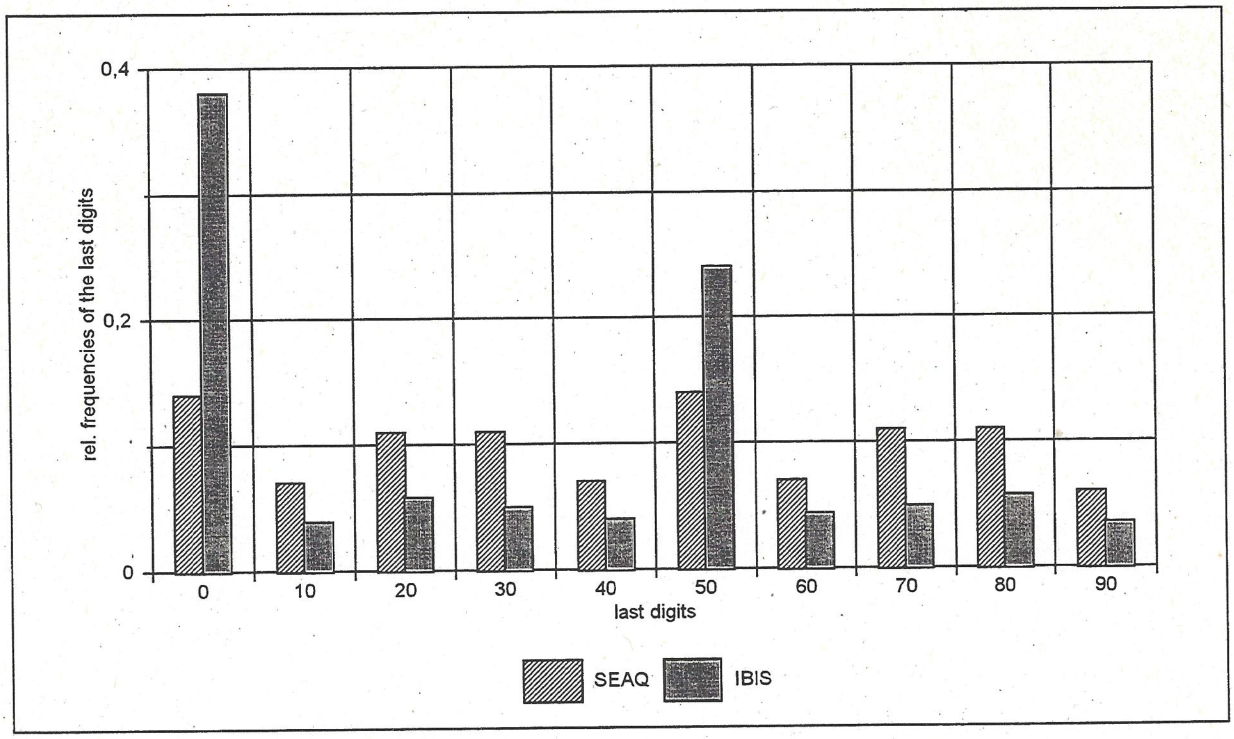
Figure 4: Relation between exactness and spread



Clustering on the market level

In this part it is looked at all quotes on a stock market. In figure 5 the result for the SEAQ (Stock Exchange Automated Quotation System; London, data from (S.J. Grossman, M.H. Miller, K.R.Cone, D.R.Fischer, D.J.Ross, 1997)) is shown together with the data of the stock market IBIS (Frankfurt, data from March/April 1993). The SEAQ data are from October 1994 and consist of 338304 inside quotes, the IBIS Data consist of 194068 inside quotes (of 30 DAX stocks). The effect of stock price clustering observed in the SEAQ data is also present in the IBIS data. However the structure seems to be different. This might be explained by a different exactness of the stocks at the different stock exchange markets. The frequencies of the last digit show structures that are similar to the bearing experiment for different levels of diffuse information and resulting exactness. The SEAQ data are finer than the IBIS data. The IBIS data correspond to a level of exactness of 5 (in figure 3) and the SEAQ data correspond to a level of exactness of 2 (in figure 3). For this comparison the exactness is not determined. This is performed in the next part in which the phenomenon is examined on the stock level.

Figure 5: Price clustering at the IBIS and SEAQ stock exchange



Clustering on the stock level

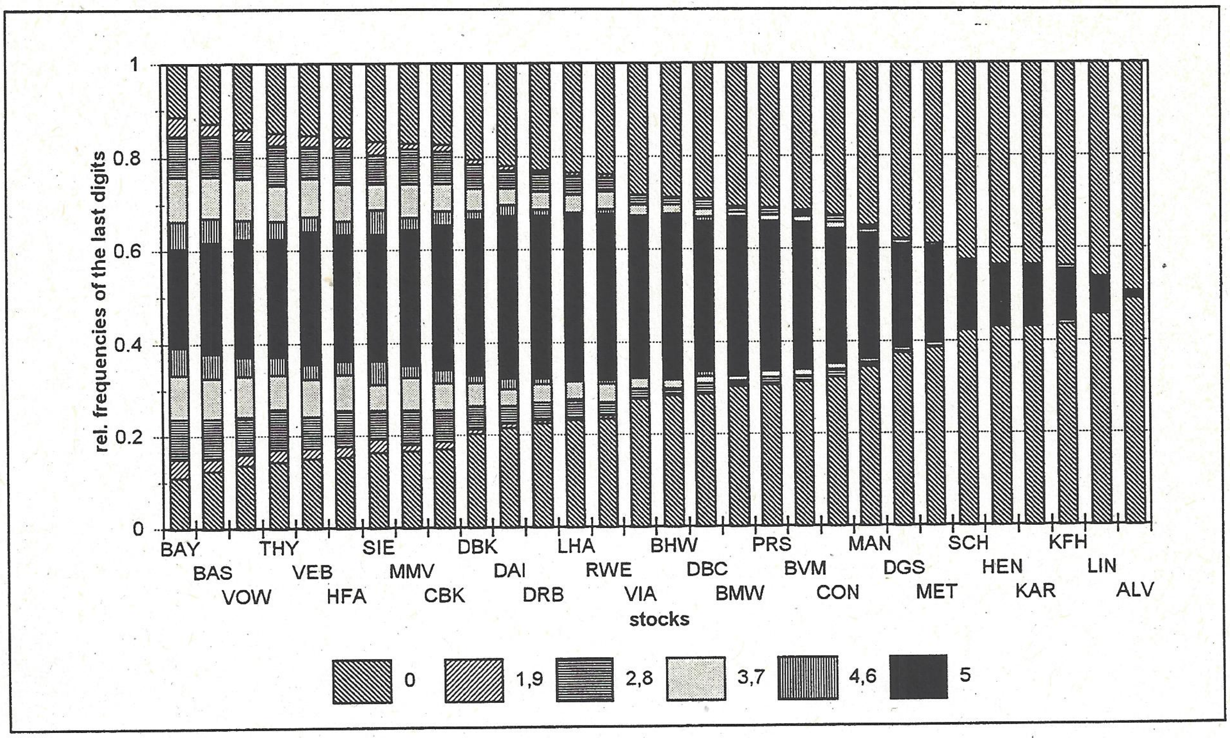
In this part further disaggregated data will be examined: 30 DAX stocks traded at the IBIS stock market. All inside quotes are considered. For all stocks the hypothesis of a uniform distribution over the last digits is rejected at the 0.01 level in a χ^2-test.

If the prices of the stocks are determined by the numerical response process the prediction for the last digits (which are multiples of 0.1 DM) is the following. For a crude exactness 0 should be the last digit with the highest frequency. By going to finer exactnesses 5 should follow, then 2,3,7,8 and finally 1,4,6,9.

Before plotting the frequencies the influence of some strategic effects were eliminated. Strategic effects of underbidding at the ask side and of overbidding at the bid side, which lead for example to higher frequencies of 9 at the ask side and 1 at the bid side, respectively, (see W. Albers, A. Uphaus, B. Vogt 1998) have been eliminated.

The diffuse numerical information is not controlled. Therefore the exactness of all prices of the stock is considered. The exactness of a stock is determined by taking all inside quotes in a period and applying the rule for the exactness of a set of data (the exactness of a set of data Γ is the crudest exactness for which it holds: at least 75% of the data have the exactness Γ or a cruder one) to this set. In figure 6 the frequencies of the last digits (which are multiples of .1 DM) are plotted versus the exactness of the stocks. On the x-axis the stocks (the abbreviations are as used at the stock market) are ordered according to their exactness. The stocks with the same exactness are ordered according to the frequencies of the prices with the crudest exactness.

Figure 6: Frequencies of the last digits of the prices of stocks versus stocks.



Looking at the maximal frequencies of the last digits first 0 is observed for a crude exactness. For a finer exactness the frequency of 5 has its maximum, 5 is followed by 2,3,7,8. Finally 1,4,6,9 occurs. This is as predicted by the numerical response process shown in figure 1.

A comparison of figure 6 with figure 3 shows that the results of the bearing experiment can be located in the plot of figure 6 starting at the maximal frequency of 5. The structures look similar.

Conclusion

The clustering phenomenon has been examined theoretically and empirically. The theoretical work is founded on the theory of prominence which describes how persons create a numerical response if they have diffuse numerical information. Applying this process to the generation of stock prices leads to predictions for the frequencies of the last digits of the stock prices. In an octal system even eighths should be more easily accessible than odd eighths. In the decimal system the prediction is: if the exactness of stock prices gets finer, the last digits of the stock prices should occur in the following order: first 0, then 5, then 2,3,7,8 and finally 1,4,6,9.

On the aggregate level the principal structures of the numerical response process can be observed. On the stock level (30 DAX stocks traded at IBIS from March/April 1993) the predictions have been examined. In these data the relation between last digit of prices and the exactness was as predicted. A similarity to a laboratory experiment in which the diffuse numerical information had been controlled was observed.

References

Albers, W. (1997): Foundations of a Theory of Prominence in the Decimal System, part I-V, IMW working papers 265, 266, 269, 270, 271

Albers, W.; Albers, G. (1983): On the Prominence Structure in the Decimal System, in Decision Making under Uncertainty, Ed. U. Schulz, Elsevier Science Publishers B.V. (North Holland), 271-287

Albers, W.; Uphaus, A.; Vogt, B. (1998): A Model of the Concession Behavior in the Sequence of Offers of the German Electrical Stock Exchange Trading Market (IBIS) Based on the Prominence Structure of the Bid-Ask Spread, IMW-working paper 287, Bielefeld

Ball, C.A.; Torous, W.N.; Tschoegl, A.E. (1985): The Degree of Price Resolution: The Case of the Gold Market", The Journal of Future Markets 5(1), p. 29-43

Christie, W.G.; Harris, J.H.; Schultz, P.H. (1994):Why Did NASDAQ Market Makers Stop Avoiding Odd-Eighth Quotes?, The Journal of Finance 49 (5), p. 1841-1860

Fechner, G.T. (1968): In Sachen der Psychophysik, Amsterdam: E.J. Bonset

Godek, P.E. (1996): Why Nasdaq Market Makers Avoid Odd-Eighth Quotes, Journal of Financial Economics 41, p. 465-474

Grossman, S.J.; Miller, M.H.; Cone, K.R.; Fischer, D.R.; Ross, D.J. (1997): Clustering and Competition in asset markets, Journal of Law and Economics 40, p. 23-61

Harris, L. (1991): Stock Price Clustering and Discreteness, The Review of Financial Studies 4(3), p. 389-415

Niederhoffer, V. (1966): A New Look at Clustering of Stock Prices, The Journal of Business 39, p. 309-313

Niederhoffer, V. (1965): Clustering of Stock Prices, Operations Reasearch 13, p.258-265

Osborne, M.F.M. (1962): Periodic Structure in the Brownian Motion of Stock Prices, Operations Research 10, p. 345-379

Stevenson, R.; Bear, R. (1970): Commodity futures: Trends or random walks?, Journal of Finance 25, p. 65-81

Vogt, B.; Albers, W. (1992): Zur Prominenzstruktur bei Zahlenangaben bei diffuser numerischer Information - Ein Experiment mit kontrolliertem Grad der Diffusität, IMW-Working paper 214, Bielefeld