Zeitschrift für Wirtschafts- u. Sozialwissenschaften (ZWS) 119 (1999), S. 1 – 22
Duncker & Humblot, Berlin

Does Information Aggregation Depend on Market Structure? Market Makers vs. Double Auction*

By Jan Pieter Krahnen and Martin Weber**

1. Introduction

Market structure and information aggregation are two important topics which have been studied extensively in the financial literature. There is a vast amount of theoretical and empirical literature comparing different types of market organization. Continuous trading systems where each trader is allowed to place bids and asks and, in addition, can accept standing bids and asks are one important form of market organization. Continuous trading at NYSE and at the German IBIS system are examples of such a trading system. An alternative form of market organization are market maker systems, where NASDAQ is an example. In a pure market maker system, two types of traders exist: market makers and ordinary traders. Market makers place bids and asks as well as they accept bids and asks. Ordinary traders can only accept standing bids and asks.[1] We will investigate four different market maker systems, altering the number of market makers between one and four. Our design thus includes the case of a single market maker (or specialist). For an overview on market micro structure, see O'Hara (1995).

Information aggregation has also been studied extensively. The key question is, if market prices correctly reflect all available information (Grossman and Stiglitz 1980 and Fama 1970, 1991). Since information can easily be controlled in the laboratory, experimental markets have proved to be very useful to study information aggregation. Starting with the work of Forsythe, Palfrey and Plott (1982), Plott and Sunder (1982, 1988) and Copeland

---

* Verantwortlicher Herausgeber / editor in charge: H. L.

** We like to thank Michael Belledin, Thorsten Freihube and especially Patrick Roll for excellent research assistance. Markus Nöth, Dirk Schiereck and Erik Theissen have helped us to run the experiments. We have benefited from comments at the DFG-Finance workshop in Liebenwalde, and workshops at the Humboldt University (Berlin), University of Bern, Wharton School (Philadelphia). The work was supported by grants Kr1221 / 3 – 1 and We 993 / 7 – 1 from the Deutsche Forschungsgemeinschaft.

[1] In the pure form considered here, ordinary traders cannot place limit orders.

and Friedman (1987, 1991), it was found that markets do aggregate information in simple cases, but they fail to aggregate when more complex settings are present.[2] Most experimental studies used a continuous trading system such as an oral or computerized double auction market as the market organization without investigating market maker systems.

Our study combines these two streams of literature and examines if the degree of information aggregation depends on the market structure. We will investigate this question by comparing double auction trading with different forms of market making, thus we will ask "Are experimental markets using double auction trading more or less efficient than those with market making and is there a difference in efficiency between different forms of market making?" An answer to this question has important implications for real world markets, i.e. for the design of optimal trading systems.

Madhavan (1992) compares (among others) a market maker (quote-driven) system with a continuous auction (order-driven) system. His results suggest that the market maker system provides greater price efficiency than the double auction which is similar to the continuous trading system he considers. The work of Pagano and Roell (1992, 1996) is somewhat related to our study as they also compare various types of auction markets with market making. However, they relate types of market micro structure to degrees of market transparency which cannot be done in our trading environments thus their results are not applicable in our setting. Recently, there has been a growing interest in experimental comparisons of market organization. Schnitzlein (1996) investigates call markets and continuous trading systems when an insider is present. Lamoureux and Schnitzlein (1995) compare a dealer market with a situation where a dealer market is competing with a bilateral search setting. Bloomfield (1996) investigates market making where two uninformed market makers face eight informed traders. However, he does not compare different institutions. Theissen (1997) compares a call market, a double auction and a market maker system on a broad spectrum of criteria ranging from market liquidity to efficiency. He finds that prices in a market maker setting convey information of high quality, but at the expense of high transaction costs. The differences in the implementation of the price process and the type of information make it difficult to compare his results to our study. In the related paper Krahnen and Weber (1997), we compare double auction and market making on the basis of liquidity and profitability.

The results of our study show that market maker systems are at least as good at aggregating information as double auction trading. This is surpris-

---

[2] See Sunder (1995) for an overview of experimental work on information aggregation.

ing in light of the efficiency properties of double auctions in much of the experimental literature.[3] However, it is less surprising in light of the theoretical results of Madhavan (1992). For different market maker systems, we find systematic variations of efficiency: systems with uninformed, competing market makers are more efficient than systems with informed, competing market makers and than systems with just one market maker. The same pattern of results shows up in belief data, which we elicit at the end of each trading period.

2. Design and Procedure

2.1 Design

To be in line with former studies on information aggregation, our design is similar to the "C" design of Plott and Sunder (1988). All subjects receive identical initial endowments and are given partial or no information about the true value of an asset. The information is different for groups of traders, but no group and no trader receives exact information about the value of the asset. The information distributed to all traders is enough for an efficient market to derive the true value of the asset.

All subjects trade one asset in one market which is organized either as a double auction system (DA) or as a market maker system (MM). The asset pays an uncertain dividend at the end of each period and afterwards it is worthless. The dividend differs between the three states X, Y, and Z, which are equally likely. Table 2.1 shows the dividends in cu (currency units). The expected value of the asset is 467 cu.

Table 2.1

Dividends depending on States

State	Dividend
X	1,000 cu
Y	300 cu
Z	100 cu

The common dividend, i.e. the true value of the asset, is determined before trading starts. One third of the subjects receive the information about one state, that will not occur (information of type 1), another third is told the second state that will not occur (information of type 2), and the last third

3 See e.g. Smith (1982).

receives no information at all. If Y, for example, is the true state, one group receives cards with the information "not X", another one cards with the information "not Z" and the third one cards containing no information, i.e. "?". In case of rational expectation, the market aggregates the information and thus knows the true value, i.e. "Y". The price will be equal (or close) to the true value (i.e. 300 cu).

A session of our market experiment runs for 13 or 14 trading rounds (periods), each period lasting six minutes. At the beginning of a period, each trader receives 3,000 cu and 6 assets. Those subjects who act as market makers receive 6,000 cu and 12 assets.[4] During trading, subjects can buy and sell assets using the trading system described below. Traders in the double auction and traders in the market maker system can short-sell up to 6 assets (market makers up to 12 assets). All agents in the market maker systems receive a credit of 10,000 cu. As described in more detail below, we run the sessions at two different locations. In Mannheim, traders in the double auction receive a credit of 3,000 credits, whereas in Frankfurt there is no credit for traders in the double auction. At the end of each period, each subject predicts the true value of the asset. Following an idea by Camerer et al (1998), traders then place either a low or a high bet on their predictions. A high bet pays 1.00 DM if the guess is correct and –1.50 DM if the guess is wrong. A low bet pays .40 DM for a correct guess and –.20 DM for a wrong guess. The expected value for the low bet is zero, if one does not know anything about the true state, i.e. for $p_{\text{true state}} = .33$. For $p_{\text{true state}} > .68$, the high bet has a higher expected value than the low bet. Finally, the true value is revealed, subjects have to give back the 3,000 cu (6,000 cu for market makers) to the experimentor, and the cash results for this period are calculated by adding up the remaining cash (or credit) and the value of the assets (number of assets times dividend). At the end of the session, the results achieved by betting are added to the final payment.

To increase the validity of our results, we run the experiment at two different universities, using three different computer programs to implement the trading systems. In Frankfurt for all sessions the program MAX is used.[5] In the double auction version, each trader can enter bids and asks or accept standing bids and asks as well as the number of assets to be traded. In the market maker version, market makers can act like the traders in the double auction. Ordinary traders, however, can only accept standing bids and asks. In Mannheim, MUDA is used as a double auction system having similar features as MAX.[6] For a market maker system, we use a system with similar

[4] As there are fewer market makers than traders and the market makers have to run the market they were given a higher flexibility through a higher initial cash and asset setting.

[5] For a description see Krahnen, Rieck and Theissen (1997).

features as MAX. As the three programs used for the simulation of the trading systems show only minor technical differences, have similar trading screens, and offer the same trading possibilities, there should be no significant influence of the trading system on the actions of the subjects.

To investigate the market maker system, we vary the number of market makers as well as the ratio of informed market makers to uninformed market makers. The acronym xMMy denotes that we have x market makers where y of the x market makers receive information. We choose the following five types of market organization (Table 2.2).

Table 2.2

Market Organization

Market Organization	Informed agents	Uninformed agents
1MM0	8 traders	1 market maker + 3 traders
4MM0	8 traders	4 market makers
4MM2	2 market makers + 6 traders	2 market makers + 2 traders
4MM4	4 market makers + 4 traders	4 traders
DA	8 traders	4 traders

The sessions are run with 12 subjects each, four of which receive information of type 1 (e.g. "not X"), four receive information of type 2 (e.g. "not Z") and four receive no information at all ("?"). In 1MM0, four traders receive information of type 1 and four information of type 2, i.e. three traders and the market maker receive no information. In 4MM0 four traders receive information of type 1, four traders receive information of type 2, and the market makers receive no information. 4MM2 has one market maker and three traders receive information of type 1 (type 2 resp.) and two traders and two market makers obtain no information. In markets 4MM4, two market makers and two traders receive information of type 1 (type 2 resp.) whereas four traders receive no information. There is one market organization (1MM0) where the market maker faces no competition in setting quotes (unlike most micro structure models, e.g. Glosten and Milgrom 1985). In 4MM0 all market makers face informed traders such as in Bloomfield

[6] We like to thank Charlie Plott for providing us with the program MUDA. Although the graphical design of MUDA differs slightly from the design of MAX, it offers the subjects identical trading possibilities and allows to exercise the same functions.

(1996). For DA, we have four traders with information of type 1, four with type 2 information, and four with no information.

Each session which uses one specific market organization, starts with two practice rounds (which are not included in the data). Afterwards, we have two periods without information, eight periods with information and again one or two periods without information.[7] We replicate each market organization: DA (six sessions), 1MM0, 4MM0 (three sessions each, thus six sessions with no market maker informed), 4MM4 (four sessions) and 4MM2 (two sessions). The double auction is replicated more often as it serves as a baseline against which all market maker markets are evaluated. Two additional sessions using 1MM0 and 4MM0 are not reported as in these first sessions the asset being traded is different.

2.2 Procedure

Subjects are graduate students of business and economics. Each session lasts about three hours and subjects are paid according to their performance in the experiment. For each subject, the cash results of the periods (not including the two practice periods) are added and divided by a number between 800 and 1,000 depending on the type of market organization. The resulting amount of money plus the money they win (or lose) from the bets is paid out in cash to each subject. The average pay out was 46.24 DM in Mannheim and 48.10 DM in Frankfurt, with a range from –1.00 DM to 121.00 DM in Mannheim and from –0.12 DM to 138.61 DM in Frankfurt. Overall, the average pay out was 47.18 DM including 1.92 DM gains from betting.

A couple of days before the sessions, students in Mannheim are trained to use the programs. In Frankfurt, they receive a longer introduction to the program before the actual session starts.[8] At the beginning of each session, students are assigned to computers and trader identification numbers by a random mechanism.

In each period, the true value is determined by drawing a ball out of an urn containing three balls labeled "X", "Y" or "Z". The drawing is public, but the label is not shown to the subjects. Afterwards, twelve cards (four cards with information type 1, four cards with information type 2, four

[7] The rounds without information are used to test the degree of risk-aversion and as comparison to the rounds with information.

[8] As we analyze the transactions of the last minute of the periods with information (round 5 – 12 of each session) the first relevant data of each session are from the last minute of the fifth round. At that point in time, the students had no operational problems with the computer-programs. We found no handling problems during rounds 5 – 12. Therefore, the different experimental training session in Mannheim and Frankfurt should not influence the results of our study.

cards showing a question mark "?") are distributed to students, such that they randomly choose a card within the restrictions given by the design (see Table 2.2). All parts of the design are common knowledge. Students are not allowed to communicate during the session except via trading.

3. Results on Information Aggregation

3.1 Market Prices: General Results

The first question we have to answer is whether or not information aggregation is observable and how good this information aggregation is. The quality of information aggregation can be derived from the difference between the market price and the true value of the asset: The closer the market price is to the true value of the asset, the better the market aggregates the information. Ideally the difference should be zero if the market price fully reflects the available information. In case of any information aggregation, the difference should be smaller for periods with information than for periods without information. To calculate the difference, we have to define what can be regarded as the market price of a period. There are quite a number of different possibilities to define this price. We can simply take the average transaction price of a six minute trading period. In case of information aggregation, subjects learn during trading, i.e. the market prices move towards the true value during the period. We consider two possibilities to capture this movement. First, we take the price of the last trade which will be denoted LP_t (last price). As the last price might be misleading due to some strategic acting, we also consider the average price of the last minute of trading, which will be denoted as LMP_t (last minute price).[9] Our analysis will mainly be based on the prices LMP_t, but LP_t prices will also be analyzed. We also estimate the equilibrium price assuming a simple time series model proposed by Camerer (1987). If the price follows the path $P_{t,i} = a + b\, P_{t,i-1}$, the equilibrium price EP_t can be calculated as $EP_t = a/(1-b)$. However, these prices turn out to be of little help. They in general are further away from the true value than both LP and LMP, i.e. the price path does not follow the underlying model.

Table 3.1 presents average last minute prices, aggregated over all types of market maker systems and over both locations in which the experiments took place (Frankfurt and Mannheim). The table indicates that the average last minute prices in periods with information are closer to the true value than in periods without information. *T*-tests for the market maker systems

[9] In case there is no transaction during the last minute of a period we take the price of the last available transaction of that period as LMP.

show that *LMP* in periods with information are significantly ($p < .01$ for each state) closer to the true value than *LMP* in periods without information for all states. Corresponding tests for the double auction are also highly significant ($p < .01$ for each state). In periods without information, i.e. periods 3, 4, 13 and 14, prices do not depend on the true value and are significantly (t-test, $p < .01$) below the expected value of 467 reflecting a certain degree of risk aversion. In periods with information, i.e. periods 5 – 12, average *LMP* are smaller for the true value 100 than those for 300 (Wilcoxon Rank Sum, $p < .01$) and smaller for 300 than those for 1,000 (Wilcoxon Rank Sum, $p < .01$).[10] Thus information is aggregated – to some extent. For each true value (state), we run a regression with *LMP* as the dependent variable, and dummies for Frankfurt and double auction as independent variables. The coefficients for the Frankfurt dummies are not significant for the extreme values 100 and 1,000. For 300 the *LMP* of Mannheim is closer to the true value than the *LMP* of Frankfurt ($p < .05$). The coefficient for the DA-dummy is not significant for 100 and 300, it is negative for 1,000 ($p < .010$) indicating that DA performs worse if the true state is 1,000. Thus as a first result we can state that market maker systems are at least as good at aggregating information as double auctions (in line with Madhavan 1992) and that we find no difference for 100 and 1,000 between the Frankfurt and Mannheim data.

Table 3.1

Average Last Minute Prices with and without Information

	Double Auction		All Market Maker Systems	
True Value (Rational Expectation)	With Information	Without Information	With Information	Without Information
100	211	371	227	403
300	317	376	325	375
1,000	489	342	613	395

In a next step, we want to investigate the degree of information aggregation. Similar to Table 3.1, Table 3.2 presents average *LMP* for all five types of market systems, with standard deviations in brackets. Average *LMP* can be misleading in case of the true value 300, as positive and negative deviations from 300 might cancel out. PAD (300) (average price absolute deviation) gives the average of absolute deviations of LMP_t from 300.

[10] We compared the distribution of market prices for true states 100 and 300, and true states 300 and 1,000. Using Wilcoxon Rank-Sum, we compare for equality of means, which is rejected at the level indicated.

Table 3.2

Average Last Minute Prices with Information

True Value	DA	1MM0	4MM0	4MM2	4MM4
100	211 (84)	232 (94)	157 (51)	190 (53)	319 (183)
300	317 (50)	325 (37)	318 (24)	305 (13)	339 (42)
PAD (300)	38	37	19	8	39
1,000	489 (143)	506 (128)	658 (235)	740 (204)	584 (188)

We see from Table 3.2, that information aggregation works quite well, as all but one *LMP* reflect information aggregation. For 4MM4 and true value 100, we get no aggregation and a high variance of prices. Variance is largest for the true value 1,000 and smallest for the true value 300, reflecting the asymmetry of possible asset values and the special role of 300 as the middle value. PAD (300) is not much larger than *LMP* (300) – 300, indicating that prices are mostly above 300 (except for DA). Table 3.2 suggests that there is a varying degree of information aggregation within the market maker systems. We will analyze these differences in more detail in Section 3.2.

Results as in Table 3.2, based on different concepts of prices, are presented in the Appendix. Table A1 gives last prices (*LP*). To gain some insight into learning during each session, Table A2 gives *LMP* of those periods in each session where the true value of 100, 300 or 1,000 appears for the last time in the session. In addition, we include Tables A3 and A4 which show the average last minute accepted asks and accepted bids. Table A1 shows that last prices mostly represent information aggregation better than last minute prices. For the true state 1,000, the *LP* of all systems are closer to the true value than the *LMP*; for the true state 100, four of the five *LP* are (slightly) better than the corresponding *LMP*, one worse. The *LMP* shows better information aggregation only for the true state 300 (in 4 out of 5 cases). As mentioned before, the prices of this state have only limited validity. It is striking that all but one standard deviations and all PAD (300) values of the *LP* are higher than the corresponding *LMP* values.

After we find that prices are (slightly) closer to the true state at the very end of a period than during the last minute of trading in a period, we will analyze if information aggregation is getting better during a session. In Table A2 we analyze the *LMP* for those periods in a session where the true state occurred the last time and compare it to the average *LMP* (Table 3.2) and average *LP* (Table A1). The pattern is more disperse: for the true value 1,000 (100) the *LMP* of last occurrence is in three (four) of the five cases closer to the true value than the *LMP* and the *LP*. Again, the results for the true value 300 are mixed. In summary it may be said that there is some learning within periods as well as within a session. We will continue our analysis with *LMP*, as this price is based on more observations but we will keep in mind that information aggregation is somewhat better for *LP*. Information aggregation is also reflected in accepted bids and asks (Tables A3 and A4). It is noteworthy that the bid-ask spreads are pretty large, which is analyzed in more detail in Krahnen and Weber (1997).

Table 3.2 presents average *LMP* and standard deviations given a specific true value (state). These values show that the distributions of *LMP* given a true value overlap. To judge information aggregation, it is also interesting to derive the true value of the asset given *LMP*. This perspective is the one of an uninformed trader who wants to infer the true value from the trading prices. We choose a simple rule to infer the true value from *LMP*:

$$
\begin{array}{rcl}
LMP < 200 & \Rightarrow & \text{true value} = 100 \\
200 \le LMP \le 650 & \Rightarrow & \text{true value} = 300 \\
650 < LMP & \Rightarrow & \text{true value} = 1{,}000.
\end{array}
$$

Note that cutoff values are equal to the conditional expected value of those subjects who receive information "not X" and "not Z". The range $200 \le LMP \le 650$ includes the expected value for subjects with no information.

Table 3.3 presents the inferred true values and shows to which extent these values are correct. In case a true value 100 and 1,000 is inferred, this inference is in all but one case correct.[11] In case a true value of 300 is inferred, in only 47 out of 105 cases this value is correct. The column "cases" shows, that 300 is inferred far too often. We see from Table 3.3, that 100 and 1,000 were only inferred in 19 and 20 out of 144 cases, i.e. well below the base rate of about 33 % each. These results demonstrate the type of error our markets perform. Market prices tend to be too close to 300. In case prices considerably move away from this middle value, they almost always correctly indicate the true state. With 75 % of correct *LMP*-inferences, the market forms 4MM0 and 4MM2 perform the best. 1MM0 and 4MM4 show the lowest degree of inference quality.[12]

Row "All*" shows the results of a second analysis using 233 and 533 as cut-off points, i.e. dividing the intervals [100, 300] and [300, 1000] into 1 / 3 and 2 / 3, to infer the true value (instead of 200 and 650). There is an increase in the number of inferences for 100 and 1,000 to 27 and 24 cases, which, except for one case, still are correct. As the data in row "All*" show, the in-

---

[11] For 4MM4 and the true value 1,000, once the state 100 is inferred.

[12] We also did the analysis in Table 3.3 for *LP* (instead of *LMP*). The corresponding Table A5 in the Appendix shows that the results are identical to the ones in Table 3.3.

ference quality increases for all market forms. We do not pursue the idea of finding optimal cut-off points as we had no ex ante hypothesis.

Table 3.3

Percentages of correct Inferences of True Values from LMP-Prices

LMP Inference	True Value	DA	1MM0	4MM0	4MM2	4MM4	Cases (in % of total)
100	100	100%	100%	100%	100%	100%	13% (19)
300	100 300 1,000	18% 51% 31%	35% 35% 30%	17% 50% 33%	22% 56% 22%	24% 36% 40%	17% (24) 33% (47) 24% (34)
1,000	1,000	100%	100%	100%	100%	83%	14% (20)
All		60%	46%	75%	75%	47%	100% (144)
All*		65%	63%	79%	81%	59%	100% (144)

3.2 Information Aggregation Depending on Market Organization

We now want to analyze in more detail, if information aggregation depends on market organization. Comparing DA and MM-systems in Section 3.1, we did not find a significant difference for true values 100 and 300, but for true value 1,000 MM-systems aggregated information weakly significantly better than DA. In this section, we consider the different MM-systems in more detail. The following analysis is based on the data for true values 100 and 1,000, as for true value 300 the *LMP*s for all market systems are about the same. Additionally, we do not consider the results of 4MM2 for this analysis as the mixture of informed and uninformed market makers in this system does not allow to distinguish between the influence factors.

We first ask whether or not competition among market makers influences information aggregation (comparing 4MM0 with 1MM0). Table 3.2 shows that 4MM0 is significantly better than 1MM0 for true value 100 (157 vs. 232, $p < .10$) and better for true value 1,000 (658 vs. 506, $p > .10$). Thus competition supports information aggregation. As shown in Krahnen and Weber (1997) in more detail, the bid-ask spread in 4MM0 is significantly smaller than in 1MM0. Thus competition leads to narrower spreads and to better information aggregation.

By comparing 4MM0 with 4MM4, we check, if information aggregation depends on the fact whether or not market makers are informed. From Table 3.2 we see that 4MM0 is better than 4MM4 for true value 100 (157 vs. 319, $p < .05$) and for true value 1,000 (658 vs. 584, $p > .10$). Thus there is evidence that uninformed market makers are better at aggregating information than a system with informed market makers. This result cannot be explained by bid-ask spreads which are slightly, not significantly larger for 4MM0 than for 4MM4 (Krahnen and Weber 1997). The result is striking for us and has to be investigated in future research. At that point, we conjecture that uninformed market makers concentrate on the information asymmetry, i.e. they are more afraid of being taken advantage of than informed market makers.

4. Results for Betting Data

Market prices are one form of data which can be used to compare information aggregation across different forms of market organizations. In addition, beliefs of traders can serve as a basis of comparison. We ask subjects at the end of each period for the true state and make them place a low or high bet on their beliefs. Similar to Section 3, we will first check if the true value can be predicted from belief data (Section 4.1) and then analyze the individual betting data (Section 4.2). The relation of market prices and betting data will be discussed in Section 5.

4.1 Betting Data: General Results

At the end of each round, we receive two pieces of information from each participant: the guess for the true value and whether a high or a low bet was chosen. To combine both types of information, we weight those guesses stronger which result in a high bet, i.e. for which the trader was relatively certain. For each period, the variable *GP* (guessing price) is calculated as follows:

$$GP = \left(\sum_{i=1}^{12} w_i G_i\right) / \sum_{i=1}^{12} w_i$$

G_i is the guess of trader i and w_i the weight of the guess which is equal to 1 (2) for a low (high) bet.[13] Table 4.1 gives the average *GP* for those periods where traders receive information, with standard deviations in brackets. For true value 300, *GPAD* gives the mean absolute deviation of *GP* from the true value.

13 We choose the factor 2, so that the high bet is as attractive as the low bet if subjects assess a probability of .68 for knowing the true state, roughly twice the base rate probability of .33. We did an additional sensitivity analysis using weights from 1.5 to 8. Increasing the weight, basically *all* average guesses decrease in value, i.e. they get closer to the true value for state 100 and further away from the true value for state 1,000. It seems as if subjects who choose the high bet, have a strong(er) tendency to bet on state 100 – independent of the true state.

Table 4.1

Average Weighted Guesses for Periods with Information

True Value	DA	1MM0	4MM0	4MM2	4MM4
100	248 (168)	258 (113)	161 (46)	211 (65)	317 (169)
300	363 (98)	355 (50)	319 (31)	302 (33)	363 (86)
GPAD (300)	87	59	26	29	73
1,000	675 (217)	698 (189)	750 (221)	818 (116)	760 (212)

Similar to Section 3.1, we see that beliefs reflect information aggregation in all but one case. Again, standard deviation is largest for 1,000 and smallest for 300. As in Section 3.1, we use the same simple rule to infer the true value given GP:

$$
\begin{array}{lcl}
GP < 200 & \Rightarrow & \text{true value} = 100 \\
200 \leq GP \leq 650 & \Rightarrow & \text{true value} = 300 \\
650 < GP & \Rightarrow & \text{true value} = 1{,}000.
\end{array}
$$

Table 4.2

Percentages of Correct Inferences of True Values from Belief Data (GP)

GP Inference	True Value	DA	1MM0	4MM0	4MM2	4MM4	Cases (in % of total)
100	100	100%	100%	100%	100%	100%	15% (21)
300	100 300 1,000	16% 65% 19%	40% 47% 13%	27% 55% 18%	25% 63% 12%	29% 53% 18%	15% (21) 33% (47) 10% (14)
1,000	1,000	100%	100%	100%	100%	100%	28% (41)
All		75%	67%	79%	81%	72%	100% (144)
All*		73%	71%	88%	88%	75%	100% (144)

Again similar to Section 3.1, we find that the true value 300 is inferred too often, even though there are more cases of 100 and 1,000 inferences through the *GP* than through the *LMP*. If the beliefs allow to infer a value other than the middle value the inference is always correct. For all market forms, the inference quality of the GP is better or equal than the inference quality of the *LMP*. Although the differences between the inference of the market forms are smaller, 4MM0 and 4MM2 still perform best and 1MM0 and 4MM4 worst. Except for the double auction, the inference quality increases further when the cut-off points 233 and 533 are used. The fact that betting data aggregates information is also reflected in the fact that in each period on average 7.90 subjects (out of 12) guess the true state, which is significantly ($p < .01$) more than the 5.33 subjects, expected without information aggregation.

4.2 Analysis on Individual Betting Data

So far, the average guesses are considered similar to the way we have analyzed the market prices. In order to analyze individual betting data, we need to know something about the homogeneity of the beliefs in each period. If all guesses were about equal (and equal to the market price), one could assume that subjects simply study the market price to derive their guesses, i.e. beliefs are really only market price plus some error. In only 17% (24 out of 144 rounds) of all periods of our experiment at least 11 (out of 12) guesses are identical. However, a definite answer about the influence of market prices on guesses is difficult as many market prices are positioned between two states.

One important question is, if informed traders give different guesses than uninformed traders. From the answer to this question we can learn, if people's private information give them an advantage compared to those people who only derive information from market trading. To check this, the hypothesis is tested if traders with private information have a smaller mean absolute deviation of guesses from true values than those without private information.[14] The data allow to reject the hypothesis that there is a significant difference between informed and uninformed traders.

In the market maker systems, we have different traders: market makers and ordinary traders. We therefore test, if market makers are better at predicting the true values than ordinary traders. Since market makers and ordinary traders obtain the same information from trading and our trading systems do not grant privileges like observation of order book etc. to any class of traders, we expect to find no difference in betting data. Indeed, the belief data shows that there is no significant difference in GP-values if one compares the group of market makers with the group of ordinary traders.

We now want to investigate if the degree of confidence (average w) is different for different classes of traders. Uninformed traders are less confident than informed traders (average w for uninformed traders = 1.38, average w for informed traders = 1.56). Although informed traders do not predict the

[14] As there were just two experiments with uninformed and informed market makers (4MM2), we did not test if there was a difference in their guessing behavior.

true outcome any better than uninformed traders, they nevertheless feel more certain about their bets. Comparing average confidence weights for market makers (average $w = 1.46$) and ordinary traders (average $w = 1.56$), we cannot come up with a conclusion. The experiment is set up in a way that the percentage of market makers being informed (45%) is smaller than the percentage of ordinary traders being informed (74%).

5. Are Beliefs More Correct than Market Prices?

Having analyzed betting and trading data, we want to investigate which of these two forms of data is closer to the true value of the asset and which form is superior to infer the true value of the asset. It is clear that both forms of data are not independent and that betting data are always gathered after trading is stopped, i.e. after all the information from trading is revealed. Based on the theory of rational expectations, a subject's betting data should diverge from market prices only if the subject has superior private information about the true value. In addition to trading data, subjects can observe bids and asks which can also be used to form a belief about the true value of the asset. Finally, market prices might be slow in adjusting to the rational expectation price, whereas beliefs already reflect this expectation. Following these arguments, betting data should be better for inferring the true value as compared to price data.

Table 3.2, Appendix Table A1, and Table 4.1 give a first comparison. They show that prices are somewhat closer to the true value than average beliefs for 100 and for 300 and further away for 1,000. In a more thorough analysis, we check for each of the 144 periods in the experiment, if *LP* – the price subjects saw before giving the guesses – is closer to the true value than *GP* controlling for the true value. Table 5.1 gives the results. The number, e.g., "83" in row "True Value 300" and column "4MM0" says that for true value 300 and market organization 4MM0 in 83% of all periods the *LP* is closer to 300 than the *GP*.

We see, that for true values of 100 and 300, market prices are closer to the true values whereas for 1,000 the averages of individual predictions outperform the market prices in 3 out of 5 cases. For the market maker systems 4MM0 and 4MM2, market prices are superior to betting data. For the true value of 300, the result is easy to explain. In this case, the LP are very close to 300 (see Table 3.2) and only one or two traders with diverging beliefs will make *GP* worse, especially due to the asymmetry of the other possible values (i.e. 100 and 1,000). For the true value of 100, *GP* is also very sensitive to individual beliefs. If, e.g., two traders believe the true state is 1,000, *GP* will be equal to 250, i.e. worse than most *LP*.

Table 5.1

Market Data give better Predictions than Betting Data: Percentages

True value	DA	1MM0	4MM0	4MM2	4MM4
100	57%	56%	67%	75%	50%
300	80%	57%	83%	60%	89%
1,000	7%	25%	56%	57%	7%

A more careful analysis of betting data and market prices is given in Figure 5.1. It shows the differences between the absolute deviations of the *GP* and the *LP* compared to the true value for all 144 rounds. The data reflect the fact that there is not much difference for state 100, the market prices are better for 300 and a strong effect in favor of guesses for 1,000.

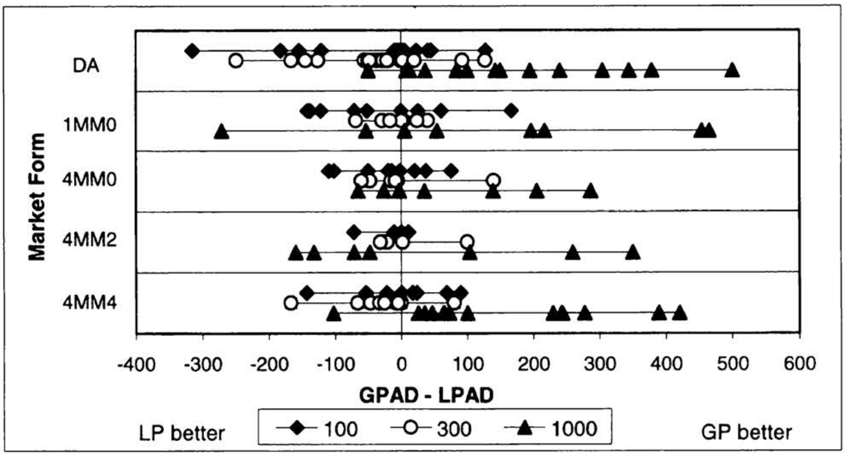

Figure 5.1: Differences between between the absolute deviations of the *GP*(*GPAD*) and the *LP*(*LPAD*) compared to the true value.

For the case of a true value of 1,000, we test if betting data derived from informed subjects is "better" than betting data derived from uninformed subjects. The idea behind this argument is that informed subjects might not have enough time to fully release their information into the market. However, this is not true: for the true value of 1,000, there is no significant difference in number of cases in which market data give better predictions than betting data for informed and uninformed subjects.

We also analyze a possible difference in inferring the true value based on market prices as compared to beliefs. For this purpose, we compare table

4.2 with the corresponding table of the *LP* (Table A5 in the Appendix). In case 100 or 1,000 is inferred, there is no significant difference as both types of data are almost always correct. Nevertheless, the *GP* infers almost twice as many times the state 1,000 as the *LP*. As this result might stem from the selection of the cut-off values, we also compare the data for the alternative cut-off values 233 and 533 but we find about the same ratio.

6. How Does Trading of Informed and Uninformed Subjects Differ?

So far, we have found as major results that markets aggregate information and that there is a difference between various forms of market organization. The ultimate question behind this result is how markets perform this aggregation, i.e. how does the conduct of an informed trader in a specific market system differ from that of an uninformed trader?

To gain more insight into which trading strategies different groups of subjects use, we study price paths during a trading round. Assume, that the current price has increased (decreased) compared to the last price and that it is getting closer to the true value. Someone who knows something about the true value should have been the driving force behind this trade. Therefore, if the price increases (decreases) towards the true value, as described above, the relative number of informed agents buying (selling) should be larger than the relative number of uninformed agents buying (selling). The data to investigate this reasoning is presented in Table 6.1.[15] When analyzing, we have to keep in mind, that in each round we have twice as many informed subjects as uninformed ones.

Table 6.1 shows the number of trades depending on price movement. Suppose the price has changed and the new price is closer to the true value. In case the price has increased, we check whether an informed or an uninformed subject has bought the asset; in case the price has decreased, we check whether an informed or uninformed subject has sold the asset. We count the trades for informed as well as for uninformed subjects.[16] Thus the numbers labeled with an asteriks in Table 6.1 say, that 99 trades toward the true value are driven by an informed agent (i.e. the informed trader bought and increased the price towards the true value or the informed trader sold and decreased the price towards the true value) and 82 trades are driven by an uninformed agent. Similarly, when prices move away from the true value, we compare if informed or uninformed traders drive the price away by buying (in case the price has increased) or selling (in case the price has decreased).

[15] All results are based on Mannheim data. Frankfurt data are nearly impossible to analyze due to the way MAX stores the data necessary for this analysis.

[16] Total trades in the last row of Table 6.1 include trades with no price movement.

Table 6.1

(x - y): Number of Trades for Informed Subjects (x) and Uninformed Subjects (y) Depending on Last Price Movement.

Number of trades	DA	1MM0	4MM0	4MM2	4MM4
A: Price toward true value	99 – 82*	134 – 23	365 – 36	153 – 54	210 – 158
Ratio	1.2	5.8	10.1	2.8	1.3
B: Price away from true value	87 – 43	51 – 17	217 – 17	74 – 17	163 – 122
Ratio	2.0	3.0	12.8	4.4	1.3
A / B	1.39	2.30	1.71	2.27	1.29
A / B Informed – A / B uninformed	1.14.–1.95	2.63 – 1.35	1.68 – 2.12	2.07 – 3.18	1.29 – 1.30
Total	589	906	1734	803	2280

In the market forms 1MM0, 4MM0, and 4MM2 informed subjects are more active than uninfomed subjects, i.e. the ratio "trades of informed subjects / trades of uninfomed subjects, i.e. row A / B" is above two. For the double auction and for the market maker system 4MM4 this ratio is below two, indicating that uninfomed subjects are on average more active than informed subjects. Additionally, these two systems have the lowest ratio between "total price movements towards true value" to "total price movements away from true value". The surprising result in Table 6.1 can be seen by comparing the ratios of informed and uninformed agents for the cases "price towards the true value" and "price away from true value", i.e. row "A / B informed – A / B uninformed". Although the ratios differ between market forms, we see no evidence that informed agents move the price more towards the true value and uninformed agents make it move away from it.

If both uninformed and informed agents drive price movements, it has to be checked if one group earns more money. We will examine this questions for ordinary traders, as in all forms of market organizations (except 4MM0) uninformed as well as informed ordinary traders are present. Table 6.2 gives the average additional profits per round in DM an informed trader made as compared to an uninformed trader. All numbers are greater than zero and the fewer market makers are informed the higher the additional profit is. These results might be attributed to the fact that there are more trades per informed agent than per uninformed agent and more trades toward the true value than away from it. The two sessions 4MM2 give evidence that informed market makers achieve higher profits than uninformed ones (.87 DM per round). For a comparison of earnings of market makers versus traders, see Krahnen and Weber (1997).

Table 6.2

Additional Profits of Informed Ordinary Traders over Uninformed Ordinary Traders (in DM).

	DA	1MM0	4MM2	4MM4
Additional Profit (in DM)	1.1	1.3	.67	.65

7. Conclusions

A lot of experimental research has investigated market efficiency using double auctions as a trading device. In our study we compare market efficiency of double auction with different market maker systems, where we vary the number of market makers and the distribution of information. We find, that market maker systems aggregate information at least as good as the double auction system. Using the ability to aggregate information as criterion for the design of optimal trading systems, we conclude that market maker systems should have no disadvantages compared to double auction systems. Second, there is a difference in performance between different market maker systems with 4MM0 and 4MM2 being superior. Third, beliefs also aggregate information and the difference in performance is also reflected in beliefs. Fourth, for two states market prices outperform belief data in predicting the final value of the assets; for one state we get the opposite relation. Finally, we cannot detect the way informed subjects bring their knowledge into prices as compared to uninformed subjects.

Clearly, this paper can only be a first step to analyze information aggregation in market maker systems. We are surprised by the difference in results between systems 4MM0 and 4MM2 and the rest. At this point we do not know why these two systems were more efficient. We think it might have to do with uninformed market makers being aware of not knowing the true state and being sensitive to informed traders' behavior. A second puzzle to be addressed in future research is that we do not find a clear difference in behavior between informed and uninformed traders, however, informed traders earn more than uninformed ones. Finally, the fact that states exist where beliefs are more efficient than market prices should be studied further.

Appendix

Table A1

Last Prices with Information

True Value	DA	1MM0	4MM0	4MM2	4MM4
100	208 (90)	228 (127)	143 (31)	193 (85)	315 (192)
300	301 (71)	344 (45)	327 (51)	328 (61)	343 (45)
PAD (300)	49	52	27	32	43
1.000	501 (177)	562 (198)	688 (247)	775 (231)	594 (216)

Table A2

Last Minute Price with Information for last Occurence

True Value	DA	1MM0	4MM0	4MM2	4MM4
100	207 (70)	224 (107)	133 (34)	171 (61)	377 (244)
300	307 (69)	336 (36)	305 (6)	315 (16)	356 (52)
PAD (300)	47	36	7	16	56
1.000	570 (168)	435 (106)	814 (222)	724 (274)	761 (125)

Table A3

Accepted Bids for last Minute

True Value	DA	1MM0	4MM0	4MM2	4MM4
100	200	191	180	120	293
300	293	273	302	285	326
1.000	405	434	635	631	530

Table A4

Accepted Asks for last Minute

True Value	DA	1MM0	4MM0	4MM2	4MM4
100	237	328	197	250	379
300	333	366	345	323	364
1.000	512	644	704	775	609

Table A5

Percentages of correct Inferences of True Values from *LP*-Prices

LP-Inference	True Value	DA	1MM0	4MM0	4MM2	4MM4	Cases (in % of total)
100	100	100%	100%	100%	100%	100%	16% (24)
300	100 300 1,000	18% 50% 32%	25% 44% 31%	9% 55% 36%	22% 56% 22%	25% 38% 38%	14% (20) 32% (46) 22% (32)
1,000	1,000	100%	100%	100%	100%	86%	15% (22)
All		58%	63%	79%	75%	50%	100% (144)
All*		56%	63%	88%	75%	59%	100% (144)

References

Bloomfield, R. (1996), Quotes, Prices and Estimates in a Laboratory Market, Journal of Finance, 51, pp. 1791 – 1808.

Camerer, C. (1987), Do Biases in Probability Judgment Matter in Markets?, American Economic Review, 77, 981 – 997.

Camerer, C. / *Nöth,* M. / *Plott,* C. / *Weber,* M. (1998), Information Aggregation in Experimental Asset Markets: Traps and Misaligned Beliefs, Working Paper, University of Mannheim.

Copeland, T. / *Friedman,* D. (1987), The Effect of Sequential Information Arrival on Asset Prices: An Experimental Study, Journal of Finance, 42, pp. 763 – 797.

Copeland, T. / *Friedman,* D. (1991), Partial Revelation of Information in Experimental Asset Markets, Journal of Finance, 46, pp. 265 – 295.

Fama, E. F. (1970), Efficient Capital Markets, Journal of Finance, 25, pp. 383 – 417.

– (1991), Efficient Capital Markets: II, Journal of Finance, 46, pp. 1575 – 1617.

Forsythe, R. / *Palfrey,* T. R. / *Plott,* C. R. (1982), Asset Valuation in an Experimental Market, Econometrica, 50, pp. 537 – 567.

Glosten, L. R. / *Milgrom,* P. R. (1985), Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders, Journal of Financial Ecobomics, 11, pp. 71 – 100.

Grossman, S. J. / *Stiglitz,* J. E. (1980), On the Impossibility of Informationaly Efficient Markets, American Economic Review, 70, pp. 393 – 408.

Krahnen, J. P. / *Rieck,* C. / *Theissen,* E. (1997), Designing an Experimental Asset Market, Working Paper, University of Frankfurt.

Krahnen, J. P. / *Weber,* M. (1997), Marketmaking in the Laboratory. Does Competition Matter?, Working Paper, University of Frankfurt.

Lamoureux, C. G. / *Schnitzlein,* C. R. (1997), When it's not the Only Game in Town: The Effect of Bilateral Search on the Quality of a Dealer Market, Journal of Finance, 52, pp. 683 – 712.

Madhavan, A. (1992), Trading Mechanisms in Securities Markets, Journal of Finance, 47, pp. 607 – 641.

Nöth, M. / *Weber,* M. (1997), Insider Detection in an Experimental Asset Market, Working Paper, University of Mannheim.

O'Hara, M. (1995), Market Microstructure Theory, Oxford.

Pagano, M. / *Roell,* A. (1992), Auction and Dealership Markets, European Economic Review, 36, pp. 613 – 623.

– (1996), Transparency and Liquidity: A Comparison of Auction and Dealer Markets with Informed Trading, Journal of Finance, 51, pp. 579 – 611.

Plott, C. R. / *Sunder,* S. (1982), Efficiency of Experimental Security Markets with Insider Information: An Application of Rational Expectations Models, Journal of Political Economy, 90, pp. 663 – 698.

– (1988), Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets, Econometrica, 56, pp. 1085 – 1118.

Schnitzlein, C. R. (1996), Call and Continuous Trading Mechanisms Under Asymmetric Information: An Experimental Investigation, Journal of Finance, 51, pp. 613 –636

Smith, V. L. (1982), Microeconomics Systems as an Experimental Science, American Economic Review, 72, pp. 923 – 955.

Sunder, S. (1995), Experimental Asset Markets: A Survey, in: Kagel, J. H. / Roth, A. E. (eds.), The Handbook of Experimental Economics, Princeton, pp. 445 – 500.

Theissen, E. (1997), Market Structure and the Aggregation of Information: An Experimental Investigation, Working Paper, University of Frankfurt.

Zusammenfassung

In dieser experimentellen Studie untersuchen wir, in welchem Ausmaß die Markteffizienz von der Marktmikrostruktur abhängt. Wir vergleichen dabei wechselseitige Auktionen mit verschiedenen Formen des Market Maker Systems. Bei den Analysen berücksichtigen wir auch die Einschätzungen der Händler, die wir am Ende jeder Periode abfragen.

Unsere Ergebnisse zeigen, daß die im Markt vorhandene Information von der wechselseitigen Auktion zu einem gewissen Teil aggregiert wird und daß das Market Maker System diese Aggregationsleistung zumindest gleich gut erbringt. Die konkrete Ausgestaltung des Market Maker Systems beeinflußt die Güte der Informationsaggregation. Diese Ergebnisse spiegeln sich auch in den Einschätzungen der Händler wider.

Abstract

In this study, we experimentally investigate to what extent market efficiency depends on market micro structure by comparing the double auction system with different market maker systems. We also elicit traders' beliefs about the true value of the asset.

Our results show that double auction markets aggregate information to some degree and that market maker systems do a job at least as good as double auctions. The type of market maker system influences the results: systems with few or no informed market makers do best in aggregating the information. The same pattern is present in the belief data.

JEL-Classification: G10, G29