

Ὀ ohne die repräsentative Stichprobe geht nicht. Da hat das Modell etwa die Bedeutung des Delphi-Orakels.¹

Ñ A prominent German pollster

Ὀ The lord whose oracle is in Delphi neither speaks out (λεγει) nor conceals (κρυπτει) gives a sign (σημαίνει).²

Ñ Heraclite²

1. Introduction

Over the last few years, experimental economists have used political stock markets to forecast elections. The first serious experiment in this area, the "Iowa Presidential Stock Market" (henceforth: IPSM), managed to predict the results of the presidential elections with an astonishing accuracy (Forsythe, Ledyard, & Shachat, 1992). Although it is the election forecast that has attracted the public's eye, this project is of secondary importance for the experimenter, whose central interest in the whole learning experiment is about information aggregation on (and through) markets. The evidence has taken the predictive success of political stock markets as a strong piece of evidence in favour of the contention that competitive markets can handle information in (almost) efficient manner, a proposition. Nor, rather, an entire bulk of propositions often associated with Nobel laureate F. A. Hayek, vis Hayek therefore labelled the "Hayek hypothesis" (Beckmann & Werding, 1994). The experiments have offered insights in

¹ From: Der Spiegel, 25/1994, 6.

² Fragment 93.

³ Cautiously enough, German pollsters reserve the term "prognosis" for the final poll conducted as voters leave the polling stations on election day. This sample excludes non-voters. No political stock markets furnish a "prognosis" in this sense because traders always have to take non-voters into account.

markets work, besides comparing favourably with opinion polls and providing teachers in economics with an excellent resource for their classrooms.

Building on their initial success in 1988, the Iowa scholars organised *Nat* times in collaboration with researchers *Na* a whole series of markets covering a host of options in different countries. Although some of them probably able to achieve as high a predictive accuracy as the original IPSM experiment had done, the accuracy of prediction was, generally speaking, quite impressive. All of these markets, however, were built on the same market institution, being organised as (computerised) double auction markets. Roughly speaking, under these rules participants may trade with each other individually through the Internet or a local network in a fashion analogous to how professional traders meet on the floor of, say, the New York Stock Exchange.

In January 1994, we set up the first *stock market* experiment with different market institutions,⁴ the 'Passau Wahlbörse' (Passau political stock market, referred to as PVW in the remainder of this paper). This experiment, comprising a market for the German Bundestag election scheduled for October 16, 1994, and another for the Bavarian State election which took place on September 25, is modelled on the Kassamarkt at the Frankfurt Stock Exchange in Germany. Again speaking rather loosely, there is no continuous trading on such a call market; instead, individual bids and asks are filed for a certain period of time, at the end of which they are treated as entered simultaneously, and the official broker, taking these orders as given, sets a market clearing price.

Our main intention in choosing this experimental setting was to put the Hayek hypothesis to a harder test than it had previously been submitted to. So far, for a variety of reasons, we held little

⁴ That we know of.

in the success of our Bavarian market as an election for given the discontinuity and comparative sluggishness of market institutions (Madhavan, 1992) one might well expect market revelation, aggregation, and dissemination of information to be impaired under call market rules—especially on a market that is run once a week. Second, experience with previous research⁵ seems to indicate that the quality of election forecasts obtained by political markets may depend on the participation of a sizable number of traders; yet, there were some 100 individuals enrolled in our Bavarian market, a few of them did actually trade. Finally, specific information bearing exclusively on the Bavarian election may well be somewhat difficult to acquire, as this local election is eclipsed by the public interest in the stream of information on, the Bundestag election.

All in all, the Bavarian market ~~is~~ ~~performed~~. Its forecasts can rival neither the subsequent double auction experiments carried out by this group, nor can it compare the accuracy typically achieved by German pollsters. While it yielded a workable prediction, considering its size and discreteness of trade, its results support the tenet that institutions matching one of our institutions—the call market institution as such, the paucity of trading days, and feasibility restrictions imposed on traders—offers and to what extent, and whether other factors such as the number of active participants play a crucial role, for central class questions that a full analysis of our market will have to answer. The second main class of questions are related to the way in which information impinges on decisions, is reflected in market prices and transmitted to traders in a call market environment.

⁵ These are, of course, not our own experiences. We owe this to Robert Forsythe and F. Nelson, who very generously let us in on some of their yet unpublished work.

⁶ For example, there are next to no public opinion polls preceding the Bavarian election, and of the few polls that are released for publication.

The present paper provides an overview of our Bavarian political call market results, along with reflections on questions that have arisen in the course of the experiment (and some tentative hints at possible answers). A more thorough analysis of the market will be presented in a subsequent paper, which will deal with our market for the Bundestag (federal) election as well as this. Hopefully, a more substantive effort might be rather long in the making, however, we elected to present a brief of our Bavarian experiences and reflections in this rather informal paper.⁷

Section 2 gives a brief overview over the design of political markets in general and the PWB call markets in particular. In section 3, we offer a few comments on the diverse measures of predictive success that we may use in evaluating our results in comparing them to rival predictions. Sections 4 through 6 contain the main results presented in this paper: Section 4 presents some demographical data about our trader population. Section 5 contrasts the official outcome of the Bavarian State election to PWB stock prices. In section 6, we describe developments on our Bavarian market, focussing on participation, turnover, the history of stock prices, and traders' behaviour. Section 7 concludes, outlining a few directions for future research.

2. PWB design principles

Experiments such as the IPISM or the PWB build on miniature markets where participants may trade claims on the contingent official outcome of a specification (Davis and Holt, 1993, 422–426). In our case, a single share in party X entitles the holder to receive a dividend of ten Deutschmarks by multiplying

⁷ This means, of course, that we are open to any suggestions our readers might put forward as to evaluating the results presented here or, in fact, the results of any call market experiment. Our address is given on the next page, and please make ample use of it.

the fraction of the popular vote received by the party X CSU managed to gain 52,8 percent of the vote in the 1998 State election, each CSU share will thus pay 5,28 Deutsche

Shares are issued in complete units consisting of one share in each of the major competing parties and one share in the party Rest of Field (RoF). Disregarding the possible rounding errors in the official election result, this unit has a final value of ten Deutschmarks and a percent corresponds to one tenth of a Deutschmark. Participants may order any number of fresh units from us or sell complete units to us at any time.

It is, however, by dealing among themselves that traders hope to make profits and risk incurring losses. The markets are organised in contrast to preceding political experiments as simple call markets corresponding to the microeconomic model of a competitive market. Traders submit bids and asks stating their reservation price and the quantity of shares to be bought or sold, respectively, in that disguise as official brokers, we collect these orders over the course of the day and use them to derive supply and demand schedules. Every Wednesday at noon, we look at the intersection of these schedules (which due to the imperfect divisibility of the commodity may contain more than just one point) to find the equilibrium price for each brand of party stock. If there happens to be a unique intersection point, such points occur at the same price, this determines the equilibrium price. If, however, the supply and demand schedules overlap vertically, we take the arithmetic

⁸ Refer to Beckmann and Werding (1994, 28–35) for an extensive description of our experiment, which also links our design to relevant experimental work. For brevity, we do not include this information.

⁹ We restricted total PWB investments to a maximum of 1000 DM.

¹⁰ From each individual trader, we accept only bids and/or asks at the best price per stock and per week. Traders may submit new bids and replace them with new ones, however.

mean of the last feasible bid and ask prices to be the equilibrium price. The maximum possible number of trades is executed at the equilibrium price,¹¹ and the markets start from scratch.

3. Evaluating predictive precision

It appears natural to judge a political stock market by its accuracy as a predictive tool, and it seems equally natural to compare this accuracy to the one achieved by public opinion polls. To do so, however, we require a workable measure of forecasting precision; if we are interested in comparing predictions as we certainly are, given the wealth of experimental data, not only across institutions, but also across elections and countries, choosing a suitable aggregate measure becomes all the more important.

A good starting point for our search is the Total Error (TE), the sum of the absolute deviations for the various parties. This implies, however, that the TE does not take account of the individual errors into account. Specifically, it cares about the number of individual errors that are aggregated not about the distribution of those errors, nor (c) their size relative to the predicted vote shares.

Problem (a) might be overcome by dividing the TE by the number of parties, thus obtaining the mean absolute error (MAE). This yields a very intuitive measure, but problems (b) and (c) remain. If you consider a prediction that is very far off track for one party (in terms of absolute distances), but fairly accurate for the others, to be worse than a prediction

¹¹ There is a rationing rule to govern the allocation when supply and demand overlap horizontally.

¹² The second row in table 1 gives the formulae for the measures we discuss here.

errors are more evenly distributed, look for a measure that penalises large errors. The shared mean error (RSME) does just that, taking account of problem (b).

Whereas the MAE and the RSME can provide an adequate summary of predictive success for polls/markets in a country, these two measures are insufficient for comparing predictions across countries that have different numbers of parties competing in elections. In other words, using measures to solve problems (a) and (b) carries a cost in problem (c). This is because dividing by the number of parties tends to even out, *ceteris paribus*, large relative errors in a set with a large number of competitors (such as German Bundestag elections). Conversely, prognoses that deal with a smaller domain fare worse, given the same TE. To see this, consider a numerical example with two countries (A and B), which differ in the number of parties involved: Table 2 shows the results for a hypothetical election and a prediction by the leading polling institutes in countries A and B, respectively. Note that both pollsters overestimate the mainstream-radical divide by the same margin (five percentage points), so that both prognoses have the same TE. The second and third rows in table 1 give the numerical values for the various measures in our simple example.¹³

¹³ In computing the numerical values for our example, we normalised the number of total votes to unity. Note that measures based on absolute errors (TE, MAE, RSME) are denominated in percentage points, whereas those based on relative errors (MRE), on the other hand, reflect

	TE	MAE	RSME	MRE	GME
	$\sum_{i=1}^N W_i - P_i $	$\frac{\sum_{i=1}^N W_i - P_i }{N}$	$\sqrt{\frac{\sum_{i=1}^N (W_i - P_i)^2}{N}}$	$\sum_{i=1}^N \frac{ W_i - P_i }{W_i N}$	$\sqrt{MAE \cdot MRE}$
A	0,05	0,0125	0,0177	0,0781	0,0313
B	0,05	0,025	0,025	0,0781	0,0442

W_i Election result for party i

P_i Price of party i stock

N Number of parties

	MainS 1	MainS 2	Radical 1	Radical 2
Election A	40	40	10	10
Election B	80	--	20	--
Prognosis A	42,5	40	10	7,5
Prognosis B	82,5	--	17,5	--

Evidently, both the MAE and RSME assert that the pollster in country A has done a better job of predicting the election than his colleague in country B, although both overestimated success of the mainstream party (-ies) by 3,125 percentage points and underestimated radical votes by 12,5%. This suggests supplanting measures based on absolute errors by their counterparts building on relative errors, for instance relative error (MRE, see table 1).

To do so, however, may well prove premature. If predicting election outcomes turns out *ceteris paribus* inherently more difficult in complex environments with

¹⁴ This problem is probably pertinent to comparative evaluation of American and German political stock markets (or polls).

multitude of parties, one might wish to allow for this gauge that gives the pollster with the harder problem headstart. Furthermore, it is clear that a large relative the more unlikely the larger the vote share of a party, respectively if we accept predictive errors distributed independently of party size. Tentatively and ignoring a host of problems crop up¹⁵, we suggest using a compound measure to be found in the last column of table 1, the geometric error (GME), which is just the geometrical mean of MAE and MRE with appropriate normalisation. In the remainder of the paper, we shall use all of the above measures on an equal footing in the final decision to the reader.

4. Some demographic characteristics of the PWB trader population

Political market traders usually fail to be representative of the general population, PWB traders are no exception.¹⁶ Since we restricted participation to university members, not what was to be expected. As table 4 indicates, our sample is very unusual even for an academic environment¹⁷, our participants turn out to be too male, too likely to be in economics, too advanced in their studies, and too right-wing

¹⁵ Not the least of which is that with this choice we come to a problem (b) again. We are, however, in two minds about this: Is it the case that a prediction that is correct for two parties must be judged inferior to a forecast with the same TE that predicted a single outcome accurately?

¹⁶ See, for instance, Forsythe and Nelson (1992, 1157). Obvious though it may seem for our economist readers that this will not, of necessity, be our market's potential predictive tool (after all, people are not known to trade according to their own political leanings), but every body's experience shows that some journal and professional pollsters find this a difficult point to grapple with.

¹⁷ We are able to correctly identify 89 out of our 100 Bavarian traders; due to missing or ambiguous trader identification, we have to exclude the remainder from the subset of Bavarian traders for the purposes of section 4.

them to be regarded as representatives even of the members of the University of Passau.

Table 3: Demographic composition of PWB participants (B only)	
Males (%)	83,2
Average age (years)	24,9
Holders of academic jobs (%)	12,4
Students: average number of semesters ¹⁸	6,2
Field of study:	
Economics (%)	71,3
Legal Studies (%)	16,0
Liberal Arts (%)	10,6
Others (%)	2,11

Beyond these demographic basics, we asked traders to indicate whether they are members of, strong supporters of, or simply sympathetic to the various parties represented on the B market. Table 4 sums up the responses of the 89 traders who managed to identify themselves. Note that a mere 28,1 % of respondents claimed to be floating voters.

¹⁸ As of entering the PWB. As the experiment spanned two semesters, this value must be taken with a grain of salt.

¹⁹ Remember that it is perfectly consistent for an individual to support for several parties, members of some party and yet weakly support another, and so on. For this reason, the percentages may add up to more than 100 % (which they do not to a certain extent).

	CSU	SPD	F.D.P.	Grüne	REP	DP	RoF
Members (%)	13,5	6,7	1,1	1,1	0	0	1,1
Strong support (%)	10,1	2,3	2,3	2,3	0	2,3	0
Weak support (%)	27,0	12,4	13,5	7,9	1,1	3,4	2,3
Total (%)	50,6	21,3	16,9	11,2	1,1	5,6	3,4

Evidently, there is a double bias in our sample. It is an economic analysis of information gathering and processing in markets that can rely neither on traders themselves nor on their level of information. However, they view their own information acquisition activities. First, around 50% of questionnaires include a series of questions as to the amount of information traders perceived to have at their disposal (on entering the market) and the primary sources of that information.

Respondents were asked to tick whether they considered themselves to be well versed in Bavarian politics (we expect an answer as 1), less well informed (2), or ill informed (3). The average point value is 1,57, which means that traders view themselves as reasonably informed about Bavarian affairs; though they believe to know a lot more about German politics in general (1,11).

Table 5 shows the sources Bavarian traders, according to their own ex ante statements, use for acquiring political information.

Used...	TV news	Radio broadcasts	Newspapers	Magazines
every day (%)	57,3	42,7	60,7	2,3
regularly (%)	15,7	19,2	23,6	39,3
sporadically (%)	14,6	19,2	11,2	38,2
not at all (%)	3,4	5,6	0	4,5
Not ticked (%)	9,0	13,5	4,5	15,7

Since ~~ex ante~~ questioning yields but a very delicate ground base one's theory on, we intend to hand out a self-questionnaire that deals with traders' information gathering habits after the Bundestag market will have closed in October.

5. Bavarian State election results, PWB stock prices, and predictive performance

How well, then, did we manage to forecast the outcome of the Bavarian State election? Table 6 compares the election to the final PWB equilibrium stock price.²⁰ Table 7 gives the values of the error measures discussed in the preceding section. Confronting PWB performance with the predictive performance of the IPSM (Forsythe, Nelson et al. and the 1992 Wahlprobe Bonn) (Schütte, 1994), an IPSM spinoff (action) experiment run for the German Bundestag election in 1990.

²⁰ Strictly speaking, the market failed to arrive at an equilibrium F.D.P. and DP stock. Supply and demand curves just did not intersect. In both cases, the prices quoted in the preceding week lie outside the week's bid-ask-spread. This is consistent with the interpretation (see section 6) that a shift in both supply and demand does reflect new information if there happens to be no equilibrium at first. Faute de mieux, we followed our ordinary procedure of taking the median of the bid-ask-spread as our prediction.

	CSU	SPD	F.D.P.	Grüne	REP	DP	RoF
El. result	52,8	30,1	2,8	6,1	3,9	2,1	2,2
PWB	47,8	30,2	4,3	7,3	5,2	2,6	3,0

Note that the experiments compared in table 7 included numbers of stock trials three for the IPSM, five for the Bonn, and seven in our case. Our reflections in section 4 are therefore, of immediate relevance here.

	TE	MAE	RSME	MRE	GME
IPSM	0,008	0,0027	0,0037	0,1443	0,020
Bonn «90	0,064	0,0128	0,0129	0,1474	0,043
PWB	0,104	0,0149	0,0252	0,2522	0,061

This is not a very pretty picture for the PWB. On the other hand, it does not permit to reject the Hayek hypothesis. The load, stock and barrel. From the repeated successes of double auction political stock markets, it tends to refer to the peculiarities of the PWB experiment. An explanation of the deficiencies of its prediction. As designed this experiment deliberately so as to reduce its predictive success. At first glance, the following spe-

²¹ We omit a comparison of PWB stock prices to forecasts by pollster, because doing so would not ameliorate the above picture substantially. It is worth, though, that out of the six (!) polls during the Bavarian election, only a series of three was released for publication to election eve. The source, *Reise*, does not belong to the small number of highly acclaimed institutions. Although its prediction turned out to be fairly accurate, neglect by our traders may be understandable.

features of our Bavarian market may have contributed to the than-perfect revelation and aggregation of information:

- ¥ the call market institution
- ¥ the small number of trading periods (one per week);
- ¥ the fact that traders were allowed to submit just a single order per market side and period;
- ¥ market thinness;
- ¥ the lack of readily accessible information on Bavarian affairs.

In section 6, we will have a closer look at some of these features of the PWB Bavarian market. All we can say here is that our results are inconsistent with a strong formulation of the Hayek hypothesis (see again Beckmann and Weisbach for a small, simple call market environment. As there are no other major stock market experiments running in parallel with our Bundestag market, the Regensburger Wahlbörse, a computerised double auction experiment based on the original Iowa design, and the Wirtschaftswoche Wahlbörse, a call market experiment run by the major German economic weekly Wirtschaftswoche²², some of these issues may be resolved once the Bundestag market has materialised.

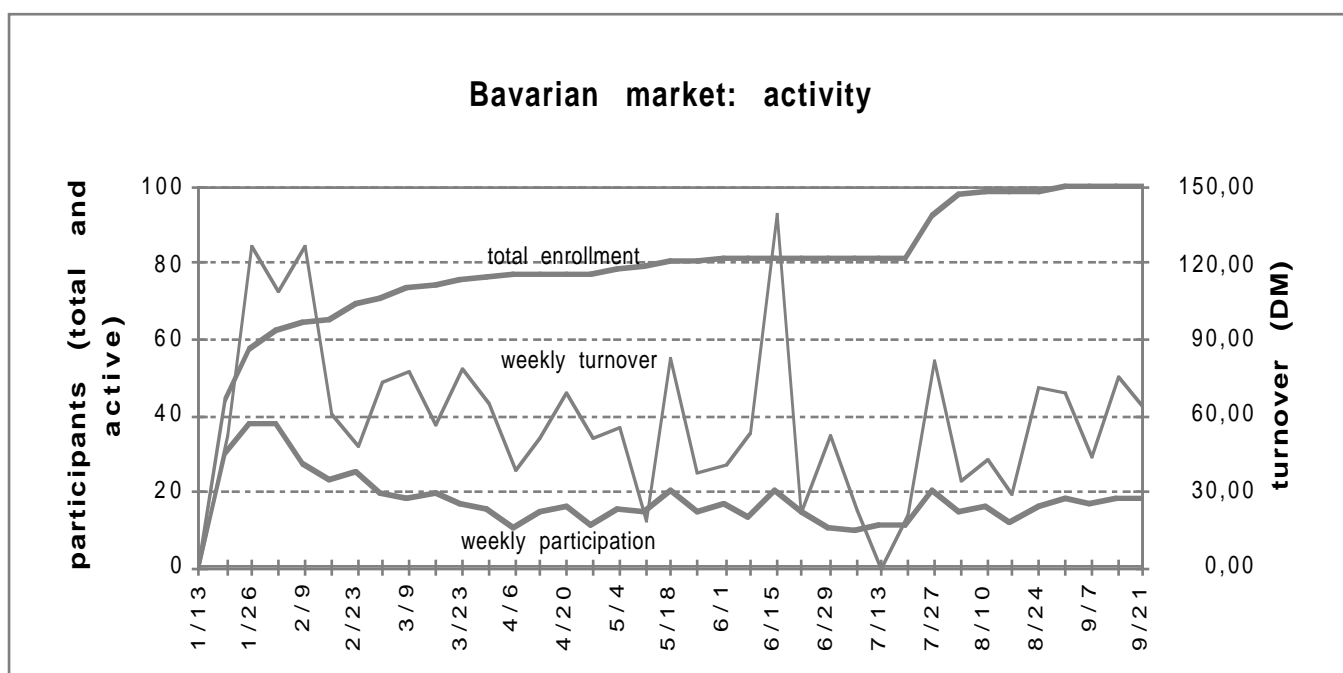
6. Trading history on the PWB Bavarian market

Our Bavarian market was clearly a small one. Figure 1 shows the development of total enrollment per trading period (the number of traders who submitted at least a single order) and the weekly turnover for the 36 weeks during which the market was

²² There are a few design aspects in which our markets diverge from the Wirtschaftswoche one. Most importantly, they do not require traders to put up any money of their own, and daily closings, have admitted a comparatively large number of traders (500) from the general public to accept any market entry later on, and offer a number of prizes to attract for the traders with the highest return. For details refer to Sch t

open. After the initial enthusiasm for our experiment out, participation quickly subsided, with less than being active in all but the first²³ six weeks.

All in all the 10 (20) most active traders²⁴ accounted for (70,1 %) of the orders we received throughout the experiment and they were involved in 45,5% (61,9 %) of their trades. A small group of eight individuals who were active in more than 50% of our trading periods.



Figure²⁵

²³ As far as participation is concerned, this puts our Bavarian experiment in the same class as the typical laboratory experiments in economics. Our experiment differs from these tests in that participants put their own money at stake; all in all, they injected 4134 Deutsche Marks into the market by purchasing a total of 297 units.

²⁴ We rank activity lexicographically, with the trading weeks in which a trader placed at least a single order as our first criterion and the number of orders submitted as second.

²⁵ Note that for the last two weeks of July, we instituted a programme under which every existing trader bringing a new participant to the market received one unit for free. This caused a surge in participation which is reflected in figure 1.

Given our Bavarian market's size, it appears remarkable that we have arrived at something resembling an election forecast. However, two aspects deserve some further interest.

According to the marginal trader hypothesis (Forsythe, Nelson et al., 1992), efficient information aggregation on a political market depends on the presence of marginal traders rather than on the sheer number of market participants. In their double auction framework, Forsythe, Nelson et al. (1992, 115) identified marginal traders as individuals placing limit orders (orders that enter the queue instead of merely accepting the best outstanding bid/offer) within a given range of the market price. Obviously, this definition can be adapted to a certain degree if we want to apply it to a call market.

Fortunately, though, the call market institution lets us supply and demand schedules in their entirety for a given period. Instead of just their extramarginal tail ends at point of time. Figure 2 shows, this opens several possibilities for demarcating the group of marginal traders:

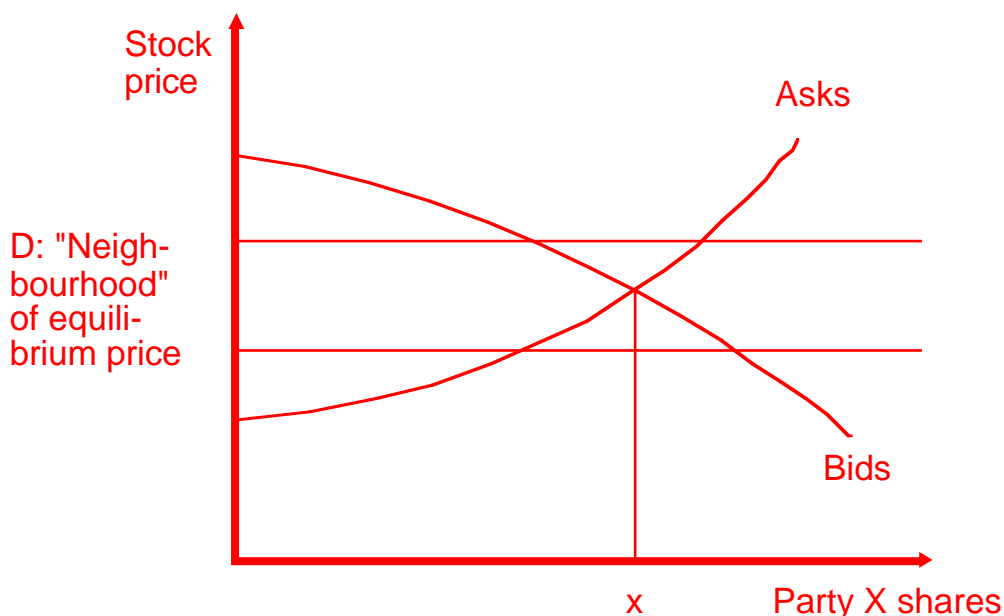


Figure 2: Ye Olde Prototypical Call Market

¥ Starting from a literal application of the theoretical definition, we might consider only orders executed at the equilibrium price as marginal.²⁶

¥ Following Forsythe, Nelson et al. (1992), we might include orders within a certain neighbourhood (D) of the equilibrium price.

¥ Finally, we might amend the last definition to include intramarginal orders within the boundaries of D those to the left of, x.

Of all these possibilities, we believe the first one to be the most attractive for several reasons. First of all, it is in keeping with the way theorists (and teachers of economics) use the term "marginal". Second, it accentuates the role of marginal orders in determining market prices, given the current state of knowledge.²⁷ Finally, and perhaps most importantly, we avoid having to choose a numerical value for D. A choice that will necessarily be fairly arbitrary. It is clear that an individual who has submitted several marginal orders to be considered as a marginal trader, and specifying the required number of such orders will introduce another arbitrary element. For the time being, though, we postpone the issue, waiting for our future results to corroborate a classification.²⁸ Whoever may have been a marginal trader, however, was probably not fully informed, as even one such trader assuming rationality and costless transactions might suffice to drive the market to a fully informed equilibrium. The presence of marginal traders in the sense of the above classification may still not be

²⁶ Because of the imperfect divisibility of PWB stock, we need a definition somewhat to include (a) the lowest (highest) price plus (b) bids or asks rationed at the equilibrium price.

²⁷ Thus, we try to draw a line between traders whose addition to demand and supply curves and those who actually close a trade at the current market price.

²⁸ It seems probable that a fair amount of marginal traders on the Bundestag market will turn out to be marginal agents on the Bundestag market.

they fail to be fully informed. Again, we need more info (such as hopefully provided by our final question: participants) to learn something definite here.

Note that trading volume and market activity (as measured by the number of bids and asks) do not appear to be perfectly synchronised. While there are periods of time where a rise in market activity is accompanied by a concomitant increase (decrease) in turnover, there are others where the opposite seems to be the case. Preliminary evaluation indicates that this might be explained by (a) the market being overvalued (or undervalued, respectively), which apparently triggered arbitrage operations, and, (b) the arrival of new information as hinted at by certain events in the political arena. This issue in the manner of (Forsythe, Nelson et al., 1992), is recurring to statistical tools, may well be due to the scarcity of observations (trading days), however.

Figures 3 and 4 depict the development of stock prices; deals with large party (i.e., CSU and SPD) stock, leaving fry for figure 4.³⁰ The characters along the market certain events in the political arena that may or may not have influenced traders' expectations. Their meaning is below. Finally, note that a hollow stock symbol denotes the absence of an equilibrium price for that stock in the respective week. There are two ways of dealing with that particular predicament (notabene, will never arise in a double auction experiment there is always a last trade to take refuge to): First of all, just carry forward last week's equilibrium price (if any); this would seem very attractive if we might take the la

²⁹ While perfect arbitrage is not possible on our markets, traders clearly took advantage of the opportunity to buy complete units from us and sell them at last week's prices (approximately) whenever they exceeded ten Deutschmarks. Quite astonishingly, the opposite strategy appeared rather less common.

³⁰ We cheat a bit by leaving out parts of the ordinate in figure 4, which exaggerates the volatility of our stock prices somewhat, but in this paper.

equilibrium price as evidence no relevant fresh information has arrived. This would, however, clearly be very unsatisfactory if the old price lay outside the current bid-ask-spread. Despite the shifting of supply and demand schedules to reflect information despite their failure to intersect, we may accept the median of the bid-ask-spread as the current prediction. In fact, this is what we did, excepting those cases where either supply or demand is missing altogether.³¹

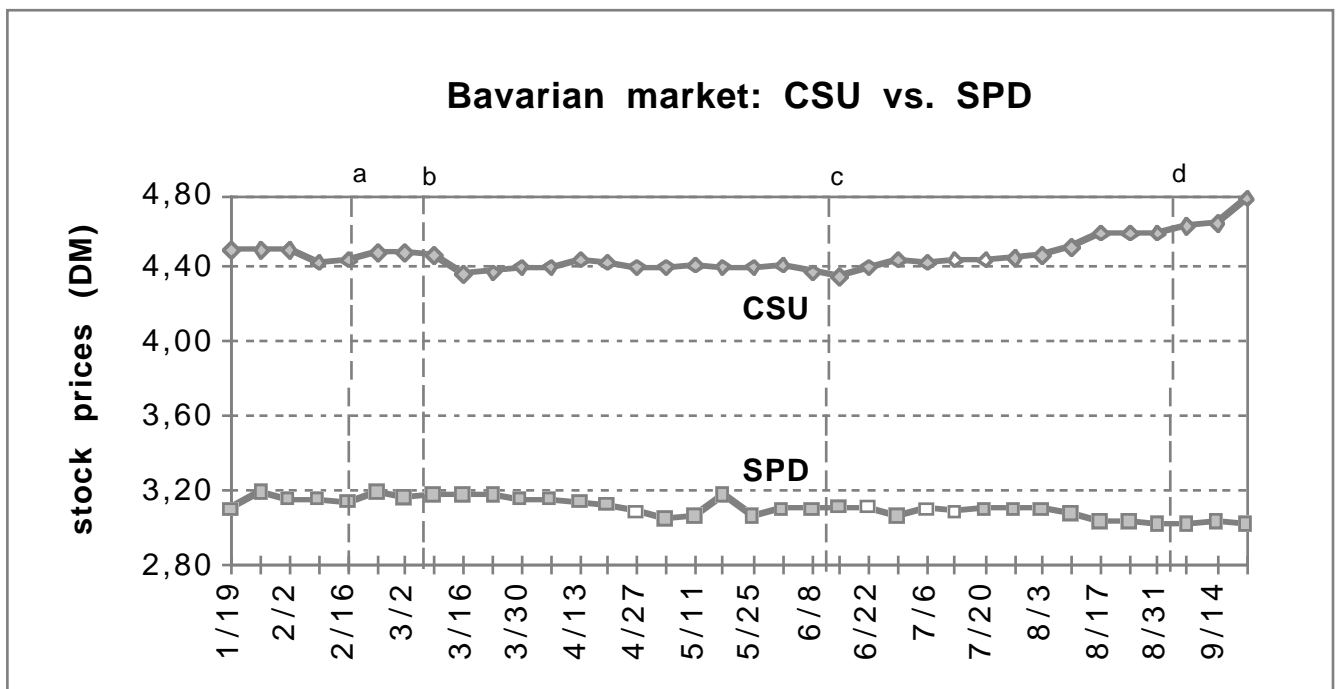


Figure 3

³¹ In these cases, we fall back to the first alternative.

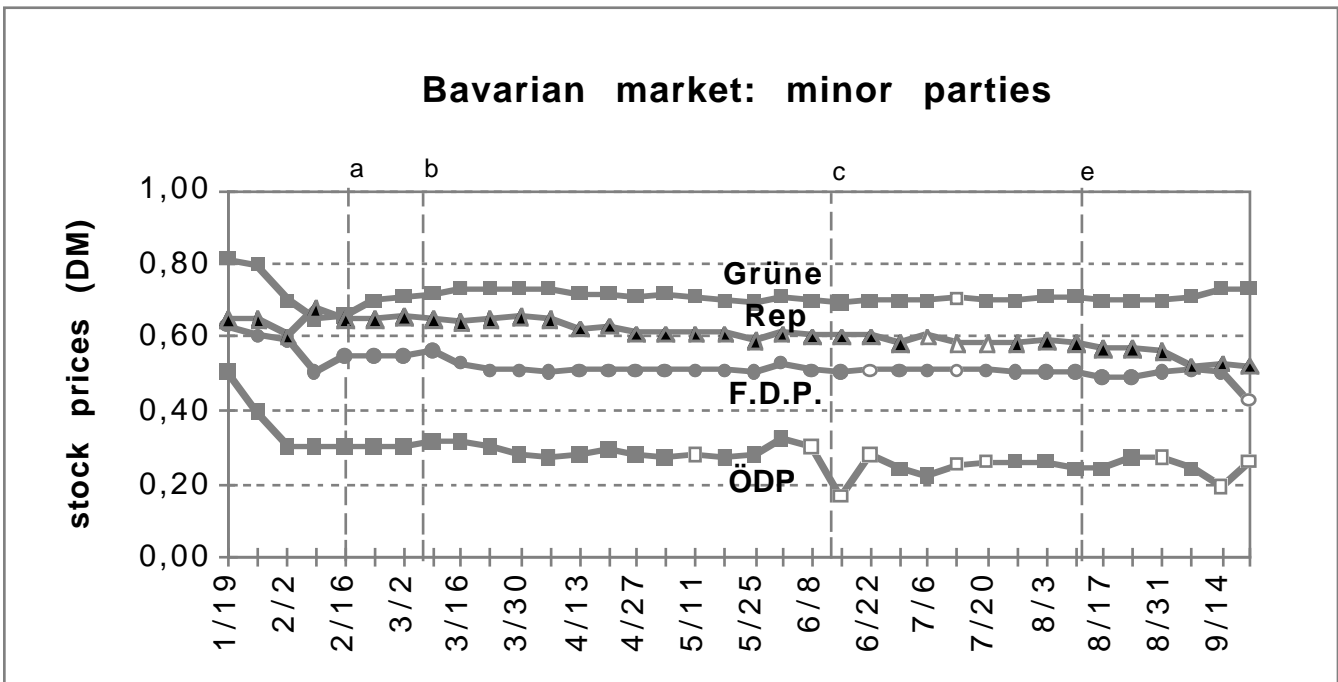


Figure 4

- a Politischer Aschermittwoch (traditional party conventions of the major Bavarian parties)
- b Local elections in some Bavarian cities and counties
- c European parliamentary elections; city council election in Munich
- d CSU party convention
- e REP chairman announces his retreat from politics and establishes connections with the radical right-wing DVU

At first glance, one remarks a striking stability of most stock prices, not an unexpected observation at all that is consistent with theoretical accounts of financial markets. After an initial period of price uneasiness (roughly speaking, after the first four weeks), prices settled down and remained fairly constant until the time of the European parliamentary election. Following this event, which a superficial inspection reveals to be the only one that appears to have influenced our Bavarian stock prices, prices generally grew more volatile and evinced a few trends: an upward one for the CSU, and a downward one for the extreme

right-wing Republikaner). Yet, most of the large number of elections that characterise the German *Superwahljahr* do not have an apparent effect on our Bavarian market. As an extreme example, consider the case of the F.D.P., whose series of Landtagswahlen prior to the Bavarian election sometimes dropping from around ten percent of the vote to mere two percent. Yet, it is only in the final four weeks we translated into some sort of commotion in the market. Price movements, however, did not go far enough, with the CSU, F.D.P. and FDP prices, although exhibiting the right direction, stopping short of the final election results.

Considering that price differentials showed significant fluctuations, it might seem plausible to put our errors in prediction (at least in part) down to the lack of a sufficient number of price signals. The lack of bids and asks reflects traders' a priori expectations based on their information gathering in the real world on the one hand and the history of political stock market prices (in part) on the other. This is why prices based on dispersed information and aggregate information may reveal more information. Observing the new price signals, however, the traders will update their beliefs. Prices do not reveal information. In the presence of noise and imperfect information, it may take quite a few of these rounds until traders reach accurate beliefs that encapsulate a sufficient amount of information.³³

A central advantage of the call market design is that it allows us to observe buyers' and sellers' surplus. In a double auction market, on the other hand, you will only observe marginal offers at any given point of time. The latter is, think of

³² In fact, this makes for an excellent laboratory that reassures us in our view that failures to arrive at an equilibrium coupled with movements of the bid-ask-spread, can be very informative indeed for subjects and experimenters alike.

³³ A careful analysis of this argument will have to spell out the learning process in some detail; this is a crucial point if we want to explain if (and how) market prices can catch up with the constant flow of fresh information, and if (and how) traders may be caught in bubbles.

total surplus in any trading period as reflecting of the information (active) traders have in this period. Unfortunately there are a host of other factors besides information discrepancies that we may adduce explaining steep and/or distant supply and demand curves. For one thing, both expect supply and demand schedules to shift leftwards as buyers (sellers) exhaust their money (stock) holdings, without any new information having arrived at all.³⁴ There are, in fact, instances on the Bavarian market where activity has petered out in this fashion. While we may hope to correct by dividing total surplus by total turnover for every period. As we have done in figure 5, other problems still remain. We do, for example, strongly suspect that strategic considerations on the part of our traders have led them to offers that differ from their true reservation prices, in turn, be only partially determined by their expectations. If this proves to be the case, the link between expectation supply/demand schedules we observe will be weak at best.

These problems arise because of a possible methodological disadvantage of our field environment. By moving the laboratory, researchers have jettisoned the complete control of all relevant parameters that can be achieved only in a laboratory, if at all. It suggests that a solution may be sought by returning to laboratory experiments that supplement political stock markets, deriving their central theoretical questions arising in the course of the field research. In such experiments one might build on a theory of strategic trader activity in markets that features diverse and asymmetric information.

³⁴ True, those agents may inject additional funds into our purchase fresh units from elsewhere, is subject to transaction costs and additional risk.

³⁵ This suspicion is based on the sort of impressionistic and casual evidence that one accumulates, albeit in an unsystematic and non-scholarly way, during the course of a long-term experiment such as the PWB.

³⁶ Judgement biases such as described by Forsythe et al., attitude effects, risk, and probably subject inexperience will also figure as explanatory factors.

For the time being, however, we return to our field experiment taking buyers' and sellers' surpluses for what they are. Figure 5 depicts the total surplus achieved on our Bavarian market as a fraction of total turnover on that market.

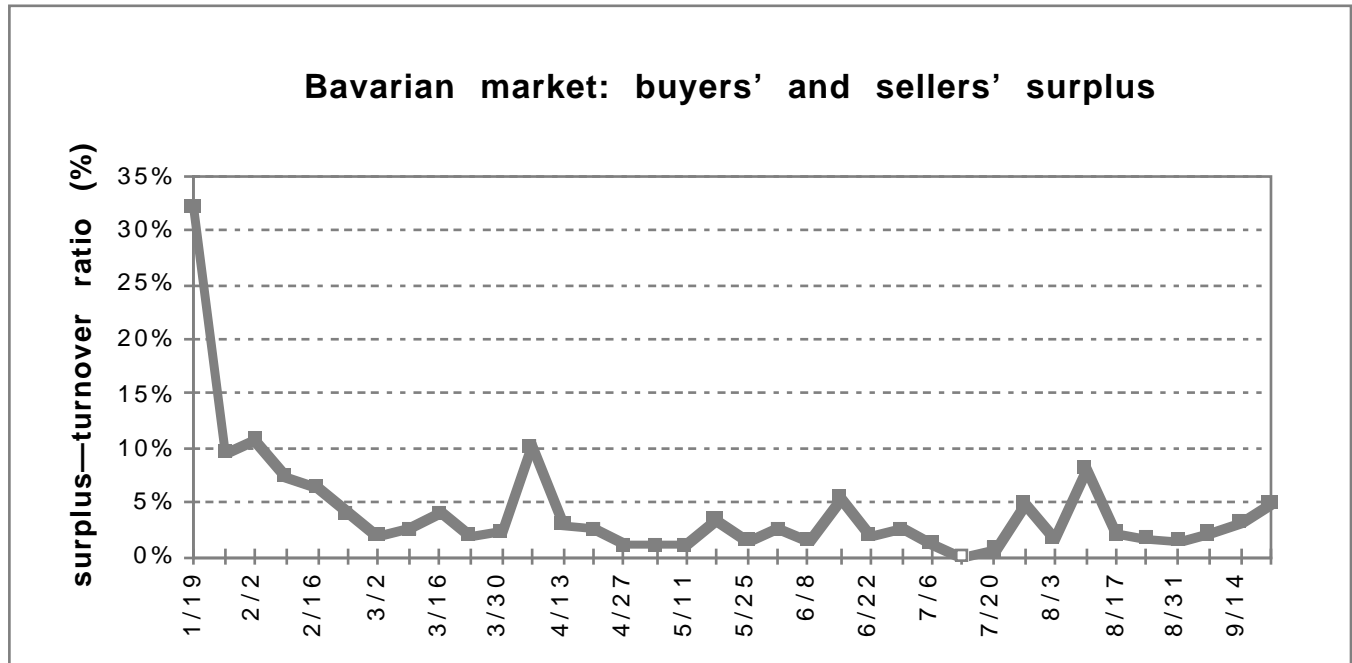


Figure 5

The first thing that strikes one as odd is the small comparative size of the surplus, which indicates that traders were quite homogenous. The development of total surplus (whether relative to turnover or not)³⁸ over time is consistent with theoretical predictions: At first, the high level of surplus because traders come to the market with heterogeneous ex ante expectations. As they update their expectations using price information, ex post expectations converge and the surplus declines. Three peaks deserve further attention: The first one (4/9) reflects a massive arbitrage operation by an

³⁷ Because of the focus of the earlier experimental literature on an auction institution, one would have to recast quite a lot of experimental work into a call market framework.

³⁸ Total surplus appears to be entirely independent of total market size. In the interest of brevity, we have not included a second graph here.

trader who offered his portfolio for sale at very low confident that market prices would turn out to be close week«s prices and thus permit to exploit market over-valuation. The second one (6/15) occurs in the week following European parliamentary election; this corroborates our impression that this event had at least some impact on expectations (conversely, events do not seem to have impressed our traders). Finally, total surplus rises again in course (and in the aftermath) of a bonus programme which led new traders into our market.

As a first step towards a careful scrutiny of information aggregation on the Bavarian market, we will have to disaggregate surplus in order to trace individual developments. This, however, will be left for subsequent work.

7. Directions for further research

It is true that the above presentation of experimental results elicit more questions than answers. If there is any justification for our putting forward so raw a picture, it is that we believe the details to be quite a long way away. While the evaluation of political stock markets built on the double auction institution rest on a huge body of experimental (laboratory) research, it is not the case for a call market. We know of no experimental research on what impact the arrival of new information will have on the behaviour of individual traders in a market. Theoretical work on that particular question is also lacking as well.

The obvious course to take is to apply the usual econometric methods to the full set of PWB data that will be available after the close

³⁹ (Kuon, 1991) has tried to link individual traders' characteristics with trading activity using the data gathered by the 1990 Wahlbörse (double auction) experiment.

our markets in October, 1994. We hope to have hinted at research strategies that might prove of value in the future. Others will surely suggest themselves to our readers.⁴ We strongly doubt that this will be sufficient to disentangle the motives from the arrival of valuable fresh information, herrings, and the obvious judgement bias that may well be ubiquitous in this kind of empirical work. To do so, one needs to supplement regressions of the common variety with an identification of marginal traders as their starting point, a game theoretic analysis of trader behaviour on a multi-period market and by laboratory experiments designed to corroborate (or to reject) the results of this theory.

References

- Beckmann, Klaus and Martin Werding (1994). *Markets and the Use of Knowledge*. Mimeo, Universität Passau: LSL 94-02.
- Davis, Douglas D. and Charles A. Holt (1992). *Experimental Economics* Princeton NJ, Princeton University Press.
- Forsythe, Robert, Forrest Nelson, George Neumann and Wright (1992). *Anatomy of an Experimental Political Market*. *American Economic Review* 82: 1142-1161.
- Kuon, Bettina (1991). *Typical Trading Behaviour in German Election Markets* Mimeo, SFB 303: B-204.
- Madhavan, Ananth (1992). *Trading Mechanisms in Security Markets*. *The Journal of Finance* 47: 607-641.
- Schütte, Christian (1994). *Wahlbörse: Kleiner Zettel*. *WirtschaftsWoche* (1994): 26-28.

⁴⁰ And we will, of course, be happy to hear of them.