

CERGE
Center for Economics Research and Graduate Education
Charles University Prague



Essays on Prediction and Betting Markets

Katarína Kálovcová

Dissertation

Prague, July, 2012

Katarína Kálovcová

Essays on Prediction and Betting Markets

Dissertation

Prague, July, 2012

Dissertation Committee

ANDREAS ORTMANN (Australian School of Business, UNSW, Sydney; chair)

PETER KATUŠČÁK (CERGE-EI, Prague)

LEVENT ÇELİK (CERGE-EI, Prague)

CHARLES R. PLOTT (Caltech, Pasadena)

Referees

PAVLO R. BLAVATSKYY (University of Innsbruck, Austria)

JORDI MCKENZIE (The University of Sydney, Australia)

Acknowledgements

Foremost, I would like to express my gratitude to my advisor Prof. Andreas Ortmann for the continuous support of my Ph.D study and research, for his enthusiasm and patience. Besides my advisor, I would like to thank the rest of my thesis committee: Peter Katuščák, Levent Çelik, and Charles R. Plott for their insightful comments.

I also thank my friends from CERGE-EI for their remarkable help with the realization of this project.

Last but not the least, I would like to thank my family and friends who were supporting me throughout writing this dissertation.

Table of Contents

Introduction	3
1 Prediction vs Betting Markets	7
1.1 Introduction	7
1.2 Definition and Examples	9
1.2.1 Prediction Markets	9
1.2.2 Betting Markets	11
1.2.3 The Success of Prediction and Betting Markets	12
1.3 Key Characteristics	14
1.3.1 Aggregation of Information	14
1.3.2 Insiders	19
1.3.3 Occurrence of Bubbles	20
1.3.4 Expected Payoff	21
1.3.5 Price	22
1.3.6 Risk-less Portfolio	22
1.3.7 Trading	24
1.3.8 Profit for the Market Maker	24
1.3.9 Liquidity	25
1.4 Payment Scheme	25
1.5 The Complexity and Flexibility of Prediction and Betting Markets	26
1.6 Conclusion	29
1.7 Appendix	30
2 Understanding the Plott-Wit-Yang Paradox	31
2.1 Introduction	31
2.2 The PWY Paradox	32
2.3 Our Experimental Betting Market	33
2.3.1 Design	34
2.3.2 Implementation	34
2.4 Results	38
2.5 Conclusion	44

3	A Prediction Market for Applications	45
3.1	Introduction	45
3.2	Literature Review	47
3.3	Our Experimental Prediction Market	49
3.3.1	Design	50
3.3.2	Implementation	54
3.4	Results	55
3.5	Optimal Market Design	70
3.6	Prediction Market Counterfactuals	71
3.7	Conclusion	73
3.8	Appendix	74
	Bibliography	79

ABSTRACT

Over the last decades, there has been a marked increase in the interest in prediction and betting markets. This interest was driven by their demonstrated capacity to aggregate information and to improve on the more traditional forecasting methods. In the first chapter, we analyze key differences between prediction and betting markets and make conjectures about the effect of these different characteristics on markets' performance. In the second chapter, we replicate the betting market experiment, cited in Plott, Wit & Yang (2003), to analyze the process of information aggregation. Based on observed market odds, Plott, Wit & Yang (2003) show that information was aggregated on their market. On the other hand, model based only on the use of private information fits traders' behavior best. In contrast to aggregate level data analysis employed in Plott, Wit & Yang (2003), we analyze individual level data and explain these paradoxical results. Finally, we conducted a CERGE-EI internal prediction market to take a well-known and experimentally tested information aggregation mechanism and implement it in a real institution for internal use to examine if it could work in this particular setting. In the third chapter, we report on the results of this market and we show that for the performance of the market, design and implementation details matter greatly, especially in the case of internal small scale prediction markets.

ABSTRAKT

V posledních desetiletích došlo k výraznému nárustu zájmu o predikční a sázkové trhy (prediction and betting markets). Tento zájem byl vyvolán zejména jejich prokázaným potenciálem agregovat informace a schopností zpřesnit využívané predikční metody. V první kapitole této dizertační práce analyzujeme hlavní rozdíly mezi predikčními a sázkovými trhy, přičemž vyslovujeme domněnky o vlivu těchto rozdílů na výkonnost trhů. V druhé kapitole se zaměřujeme na proces agregace informací pomocí replikace experimentálního sázkového trhu uvedeného v publikaci Plott, Wit & Yang (2003). Plott, Wit & Yang (2003) prostřednictvím kurzů na jejich sázkovém trhu ukazují, že informace byly agregovány. Paradoxem je, že model založený pouze na používání soukromých informací vystihuje chování jednotlivých obchodníků nejlépe. Na rozdíl od analýzy agregátních dat, kterou zvolili Plott, Wit & Yang (2003), my analyzujeme data na úrovni jednotlivých obchodníků a vysvětlujeme tento paradox. V poslední kapitole uvádíme výsledky našeho interního predikčního trhu zorganizovaného na institutu CERGE-EI. Tento trh nám umožnil testovat známý, experimentálně ověřený mechanismus pro reálné vnitřní potřeby této instituce. V rámci tohoto experimentu jsme poukázali na fakt, že design a implementace predikčního trhu jsou nesmírně důležité pro jeho správné fungování a to obzvláště, když se jedná o interní trh s nízkým počtem obchodníků.

Introduction

The trustworthiness of popular judgements and the reliability of collective decision making have been of substantial interest for a long time. A hundred years ago, British scientist Francis Galton attended a weight judging competition at Plymouth. Almost eight hundred townspeople estimated the weight of an ox. Some of the participants were experts at judging the weight of livestock (butchers, farmers) while others were guided only by their fancies with no insider knowledge of cattle. Surprisingly enough for Galton, the mean estimate, which can be interpreted as the collective wisdom of the Plymouth crowd, was pretty accurate.

Galton's experience suggests that under some circumstances, groups are remarkably intelligent and even smarter than the smartest people in them. Galton saw much more in this competition than just popular entertainment. He points out that the average competitor is probably as well suited for making a just estimation of the ox, as "an average voter is of judging the merits of most political issues on which he votes" (Galton 1907, p.450); could we continue "as an average employee of estimating the level of sales" or "as an average citizen of predicting the success of a certain public policy"?

Consider a firm that is about to make a significant investment in the launching of a new product on the market or a risk-manager whose long term trades are based on

the forecasts of macroeconomic indicators. In these scenarios a high accuracy of forecasts is of the utmost importance. Predicting the outcome of uncertain events like sales forecasts, macroeconomic indicators, political or sporting events clearly requires that we have relevant and sufficient information with which to make our prediction as accurate as possible. Yet, as Hayek (1945) (p.519) puts it, "the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form, but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess."

The first experimental evidence of a markets'ability to coordinate information to reveal an equilibrium competitive price can be found, for example, in Smith (1962), and studies explicitly designed to analyze the ability of different market mechanisms to aggregate information include, for example, Plott and Sunder (1982) or Plott and Sunder (1988).

Over the last decades, there has been a marked increase in the interest in continuous double auction markets and call markets (prediction markets) and parimutuel and fixed-odds betting systems (betting markets). This interest was driven by their demonstrated capacity in some settings to aggregate information and to improve on the more traditional forecasting and decision-making methods.

In this experimental dissertation, we focus on the process of information aggregation in prediction and betting markets, which is well documented (Plott, Wit, and Yang 2003) but not so well understood (Plott, Wit, and Yang 2003; Roust and Plott 2005; Berg, Forsythe, Nelson, and Rietz 2008). We developed software for our experiments and conducted three experimental markets. We analyze data from these experiments and together with summarizing and classifying key features of both prediction and betting markets, we try to shed some light on why prediction and betting markets are successful and how various design and implementation features affect their performance.

Chapter 1: Prediction vs Betting Markets. The results from both prediction and betting markets are mixed. While there is extensive evidence that these markets provide unbiased estimates in various settings (e.g. Debnath 2003 or Berg et al. 2008), there are

also known cases of not-so-successful markets (e.g. Berlemann, Dimitrova, and Nenovsky 2006 or Jacobsen, Potters, Schram, van Winden, and Wit 2000). These mixed results together with our internal prediction market project (reported in the final chapter) led to a review of the key characteristics of prediction and betting markets presented in the opening chapter of this dissertation.

The hope was to provide some guidelines for a proper choice of a market mechanism depending on the particular application. We observe that prediction markets are extensively used in many different areas while the use of betting systems is concentrated on the predicting outcomes of sport events.

We conjecture that this is due to several advantages of prediction markets over betting markets. In particular, prediction markets are more efficient and flexible and hence, for most purposes, a more suitable information aggregation mechanism compared to simple but rather rigid betting markets.

Chapter 2: Understanding the Plott-Wit-Yang Paradox. The main focus of this chapter is on the process of information aggregation in prediction and betting markets. In particular, in the second chapter of the dissertation, we report the results of a parimutuel betting market based on Plott, Wit, and Yang (2003). We focus on traders' behavior in terms of acting on private signals versus updating one's beliefs based on observed behavior of other participants.

We find extensive support for our hypothesis that participants in our betting market learn from other participants and form a weighted average of their private signal and of the market's public signal. In the early stages of our experiment, participants tend to trust their private signals more. However, as the experiment progresses, participants learn that the public signal is more informative and increase the weight put on this source of information.

Chapter 3: A Prediction Market for Applications. In the third chapter, we test the aggregation of information in a specific setting. We conducted two internal experimental

prediction markets at CERGE-EI, Prague, to forecast the number of applications received for the 2009/2010 Ph.D. program. The results of this project are promising.

While we found certain inefficiencies (the sum of prices is greater than one, not winning events are overpriced), overall our short-term prediction market worked very well. The success of our long-term prediction market was not so overwhelming, but we conjecture that the rather poor performance can be explained by the incentive system used.

Our experimental markets show that while prediction markets have the potential to serve as a decision making tool, we have to be careful about their design and implementation especially in the case of internal, small-scale prediction markets.

Chapter 1

Prediction vs Betting Markets¹

1.1 Introduction

The popularity of prediction and betting markets has increased significantly over recent years. This interest was driven by their ability to aggregate information and to increase the accuracy of predictions compared to more traditional methods. In order to make our prediction as accurate as possible, we need to have all relevant information which is usually possessed by a large number of individuals and rather difficult to gather. Prediction and betting markets seem to have a potential to aggregate all the information available and therefore help make reliable predictions of various uncertain events: be it political or sporting events, sales forecasts, or macroeconomic indicators.

The terms *prediction* and *betting* markets are often used to describe several different market mechanisms. The most common² are the:

- pari-mutuel betting market
- fixed-odds betting markets
- continuous double auction (prediction market)
- call market (prediction market)

¹**All errors remaining in this text are the responsibility of the author.**

²This list of market mechanisms is not complete, we only mention the most wide-spread types of market. There are other possibilities, such as point spreads offered by bookmakers. However, these types are not relevant for the purpose of this chapter.

In this chapter, we analyze two types of markets: *the pari-mutuel betting market* and *the call market*. There are two reasons for this choice.

The first and the main reason comes from the objective of this dissertation. In all its chapters, we concentrate on internal prediction markets used within corporations or non-profit institutions. These markets are usually small and organized to acquire the prediction of the event of interest (e.g. the end date of a project, level of sales, etc.) rather than to make a monetary profit. It might be the case, that the market organizer either does not have sufficient information to set reasonable market odds, or he does not want to set market odds and influence traders by suggesting what the "correct" odds are (as needed in fixed-odds betting markets often used in sports).

Further, a small number of participants usually leads to a relatively low level of activity on the market. If the market is open continuously (like in a continuous double auction market), participants might be discouraged by a low activity throughout the day and lose interest in participation.

Therefore the likely candidates for an internal prediction market are a:

- *call market*, where limit orders are matched and the market cleared once or twice a day
- *pari-mutuel betting market*, where the market odds are not set by the market maker but result from the wagering of all traders

The second reason for choosing the pari-mutuel betting market and the call market is the fact that in this experimental dissertation, we report the results from two experiments that we conducted. The first experimental market was organized as a pari-mutuel betting market (inspired by Plott, Wit, and Yang 2003) and is analyzed in the second chapter of this dissertation. The second experiment was a call market organized at CERGE-EI, Prague and is described in the third chapter. After conducting these two experiments, natural questions arose: What are the major differences between these two markets, and how do their key characteristics affect their performance? Is one mechanism more appropriate than the other (in some situations)?

To our best knowledge, there is no research which would provide suggestions for the proper choice of a market mechanism. This chapter therefore tries to fill this void. We list and analyze the key characteristics of both types of markets and make conjectures about the effect of these key characteristics on their performance. Our analysis is mostly based on existing experimental and theoretical literature. In other words, we try to summarize the known facts about both types of markets and discuss both their advantages and drawbacks. We also provide several conjectures which are based on our experience and the results of experimental markets discussed in chapters 2 and 3. This chapter may be found useful by experimenters before they make a decision about market mechanisms.

1.2 Definition and Examples

In this section, we explain how both a pari-mutuel betting market and a call market work, and we provide some examples of areas where these markets are typically used. In order to make our arguments and conjectures more general, for the remaining part of this chapter we will talk about prediction and betting markets in general. Most of the key characteristics that we will mention hold for all types of prediction markets (continuous double auction and call markets) or betting markets (fixed-odds and pari-mutuel betting markets). If this is not the case, we will make an explicit note about their differences.

1.2.1 Prediction Markets

A continuous double auction is a type of a prediction market, where potential buyers submit their bids and potential sellers simultaneously submit their ask prices to an auctioneer, and then an auctioneer chooses a price p that clears the market: all the sellers who asked less than p sell and all buyers who bid more than p buy at this price p .

A call market is a market in which trading in individual securities occurs at specific times as opposed to continuously. In a call market, all orders to buy and sell a particular security are assembled at one time in order to determine a price at which most of the orders can be executed. Call markets are frequently used in situations in which there are

few securities and participants. Because the call market groups transactions together, there is a substantial increase in liquidity.

The sole difference between a continuous double auction (used for example in Berg, Nelson, and Rietz 2003) and a call market (for example the CERGE-EI internal prediction market analyzed in chapter 3) is the way of matching bids (offers to buy an asset for at most a given price) and asks (offer to sell for at least a given price). In the continuous double auction market, bids and asks are matched continuously. Moreover, best current bids and asks are listed and can be accepted by traders immediately. In a call market, limit orders are made privately by traders and with a pre-determined frequency, these orders are matched and trades are made (if bid prices exceed ask prices).

Examples of *prediction markets* include, for instance, the famous Iowa Electronic Market, which has been used to predict the results of American presidential elections since 1988. In this market, participants can trade with the assets that pay \$1 if the next president is a Democrat and \$0 otherwise. The market price of this asset, for example 80 cents, is then interpreted as 80% probability that the next president will be a Democrat. Similarly, participants can trade with assets whose payoff is tied to the next president being a Republican. In this particular Iowa Electronic market, participants can trade with two mentioned assets. In general, there can be multiple assets on the market, and they are chosen in such a way that they cover all possible outcomes of a predicted uncertain future event.

The Iowa Electronic market is analyzed by Berg, Forsythe, Nelson, and Rietz (2008) who summarize a dozen years of election futures markets research and find that the election-eve market forecasts generally predict better than do the major national polls. Berg, Nelson, and Rietz (2003) point out that there has been no analysis of prediction markets' long run predictive power, and they are the first to show that markets are generally much better predictors than are polls months before the elections.

Further examples of political prediction markets include, for example Jacobsen, Potters, Schram, van Winden, and Wit (2000) who report the results of their political pre-

diction market in the Netherlands. The authors find that their markets were not as accurate as the Iowa prediction markets. Despite quite a large interest and a high number of traders, their markets failed to outperform the national polls. Another example is the first experimental market in the Czech Republic designed to predict the outcome of parliamentary elections in 2002 (see Cahlik, Geršl, Hlaváček, and Berlemann 2005).

Apart from being employed in predicting the outcomes of political events, prediction markets appear to be useful decision-making tools for corporations. Hewlett-Packard, for example, pioneered applications to sales forecasting and, due to their success, has since integrated prediction markets in several business units (Chen and Plott 2002).

The experimental prediction market created at Siemens, which aimed to estimate the expected delay in completing a certain project, showed satisfying forecasting potential (results are summarized in Ortner 1997 and Ortner 1998). Such cases suggest that prediction markets used in corporate level decision-making processes, e.g. as an indicator of the probability of the success of a project, indeed show promise.

Apart from probably the most well-known Iowa Electronic Market for predicting election outcomes, other examples are mostly web-based markets with the possibility to bet on sports, entertainment, or economic indicators and their turnover ranges in hundreds of thousands to millions of dollars (for example intrade.com).

1.2.2 Betting Markets

The second type of a market, a **pari-mutuel betting market**, or simply *a betting market*, as typically used in horse racing and other sports events, is based on a betting system in which all bets are put together and the payoff (odds) is determined by dividing the total amount of money invested by the amount betted on the winning horse.³

Pari-mutuel betting differs from fixed-odds betting in that the final payout is not determined until the pool is closed. In **fixed-odds betting**, the payout is agreed at the time the bet is sold. Odds can change during the existence of a betting market — for

³The amount of money available for a distribution to the winners is usually lowered by the house charge.

example, if a market maker realizes that he missed some important fact and needs to readjust the odds — but for the traders, only the odds under which they place a bet matter.

In this chapter, the use of "betting markets" refers to pari-mutuel betting systems because we concentrate on small-scale, internal prediction markets. On these internal markets, there is no market maker who would set the market odds. On the contrary, the hope is to use this mechanism to aggregate dispersed pieces of information and find out what the proper odds are, which are interpreted as the likelihood of a given event. Pari-mutuel betting systems work without the need of setting the market odds centrally, and there is evidence that pari-mutuel betting markets provide unbiased forecasts of predicted outcomes (see, for example, Debnath 2003 or Plott, Wit, and Yang 2003).

As mentioned for example in Sauer (1998), the pari-mutuel betting system is used by racetracks in North America, France, Hong Kong, Japan, Australia, and Great Britain. Fixed-odds betting is very common in sports betting (betfair.com, bwin.com, etc.). A growing number of companies organize usually web-based markets with both a pari-mutuel betting system — mostly for horse racing and fixed-odds betting markets — for other types of sports (for example centrebet.com). For more examples of web-based markets (both prediction and betting mechanisms) see, for instance, Wolfers and Zitzewitz (2004) or the Appendix of this chapter.

1.2.3 The Success of Prediction and Betting Markets

Evidence (empirical and experimental, see Berg et al. 2008 or Chen and Plott 2002 as an example) so far suggests that both prediction and betting markets are surprisingly accurate in predicting a wide range of uncertain future events: be it sports events, election outcomes, or the completion dates of corporate projects. The reasons seem to be manifold: First, they are often able to aggregate information that is dispersed (Chen and Plott 2002); second, anonymous trading makes participants more likely to reveal what they really know (Berg, Nelson, and Rietz 2003); and, finally, well-constructed prediction markets are difficult to manipulate (Rhode and Strumpf 2004).

We argued above, that there is no doubt that both prediction and betting markets work well. However, it is not so clear whether there are some apparent advantages of one mechanism over the other. In other words, do prediction markets perform systematically better than betting markets for example? The answer to this question appears to be negative. If one mechanism or the other was better in all circumstances, there would not be a need for both of them. This brings us to an even more interesting question. Do prediction markets prevail in some settings and betting markets in others? Are there environments for which one or the other type of market is more suitable? To our knowledge, no research provides insight into these issues. The authors of prediction and betting markets literature provide no justification for their choice of market.

In this chapter, we analyze the key characteristics of both prediction and betting markets. Whenever possible, we support our arguments and conjectures with all relevant theoretical and experimental articles on prediction and betting markets that we know of.

We exclude empirical research on prediction and betting markets because this area of research is focused on understanding traders' behavior and its consequences (efficiency, favorite long-shot bias, etc.) rather than on decisions about the basic market design such as the choice between a prediction and betting mechanism.

Based on the key characteristics of prediction and betting markets, we try to understand what the proper choice of the market might be (if there is one) in various circumstances. We conjecture that these key characteristics have implications in terms of market simplicity and flexibility combined with the capability to provide accurate predictions (efficiency), and this determines the suitable form of the market in various settings.

While a betting market seems to be a proper choice in large-scale markets with possibly unexperienced traders (e.g. sports betting), a prediction market appears to be more suitable in small-scale markets with mostly experienced traders and high expectations about prediction accuracy.

1.3 Key Characteristics

Both prediction and betting markets generate predictions of outcomes of uncertain future events. We summarize their key characteristics⁴ in the table below, and we discuss every characteristic in the following sections.

	Prediction Market	Betting Market
Aggregation of Information	yes	yes
Insiders	enter early	enter late
Occurrence of Bubbles	lower	higher
Expected Payoff	known	known only ex-post
Price	market clearing	fixed
Risk-less Portfolio	exists	does not exist
Trading	continuous	no re-selling
Profit for the Market Maker	positive	positive
Liquidity	limited	unlimited

Figure 1. A summary of key characteristics of prediction and betting markets.

1.3.1 Aggregation of Information

In this section, we establish the fact, that both prediction and betting markets have the capacity to aggregate information. By information aggregation we mean the fact, that the likelihood of individual events determined by market prices is close to the probability distribution that would exist if all traders pooled their private bits of information.

Moreover, we point out that both types of markets have sufficient capacity to attract traders with accurate information about some uncertain future events. We refer to these traders as insiders. There is a very close relationship between the presence of insiders and the accuracy of predictions. In the subsequent section, we discuss which market mechanism gives insiders more power to drive market prices towards their true value and hence leads to more accurate predictions.

Although it is already well established that prediction and betting markets can be used to aggregate disseminated pieces of information and can generate accurate predictions of

⁴Some of these characteristics are supported by existing literature, others are our conjectures yet to be tested. Which is the case is stated when we discuss these characteristics in further detail below.

a variety of uncertain future events such as elections, project completion or sports events, it is not well understood why these markets perform so well.

Notwithstanding a growing body of prediction and betting markets literature, theoretical questions concerning how traders learn information from the market price (odds) and how markets aggregate dispersed pieces of information remain unresolved. Plott, Wit, and Yang (2003) state that there is no clear theoretical reason why *pari-mutuel systems* should aggregate information at all. However, the implicit prices on their experimental markets are very close to the prices that would exist if all agents pooled their information and made decisions on the basis of the pooled data. This observation suggests that the information in their markets *does* aggregate.

Relatedly, Roust and Plott (2005) raise the important question whether the odds that emerge from a betting process have the characteristics of the probability distribution over the possible states that results from the pooling of information held by all individuals. Plott, Wit, and Yang (2003) suggest that the answer is positive except for the fact that the odds frequently overstate the probability of rare events and underestimate the probability of favorites. Further, Berg and Rietz (in Hahn and Tetlock 2006, pp 142-169) and Forsythe et al. (1992) present open issues that need to be addressed in future research including a theoretical model of prediction markets consistent with observed trader behavior.

Berlemann and Nelson (2005) also present an interesting open theoretical question concerning the way in which traders learn information from the market price. This question is a key problem in understanding the process of aggregating information on *prediction markets* because the market price is not a simple average of all traders' prior expectations. As traders enter the market, they learn the expectations of other participants at least to some extent and can update their subjective predictions.

This process can be of major importance especially in the case of small-scale prediction markets implemented within a firm where the actions of one trader have a relatively big impact on the market price. A full understanding of the process of aggregating information in the market is of great importance for the decision maker who can regulate

the liquidity and the exchange of information in the market in order to enhance its effectiveness. In other words, the aggregation of information in prediction markets is closely related to the optimal role of the market maker.

Ottaviani and Sorensen (2006a) model trade and price formation in a *prediction market* with risk-averse individuals having heterogenous prior beliefs, heterogenous private information, and heterogenous attitudes toward risk. They show that the aggregation of information cannot be easily separated from an aggregation of beliefs.⁵ The authors posit that the market is in a fully revealing rational expectations equilibrium, i.e. traders make correct inferences from the prices, given common knowledge of the information structure and prior beliefs. In other words, the authors assume that prior beliefs, signal distribution, and the rationality of all traders are common knowledge; therefore, they, in fact, assume that information does aggregate and do not model it explicitly. The authors' results suggest that the process of information aggregation can be modeled assuming common prior beliefs and heterogenous private information about the probability of each outcome.

Plott, Wit, and Yang (2003) present several theoretical measures of information aggregation on *pari-mutuel betting markets*. Further, they experimentally test the ability of pari-mutuel betting markets to work as an information aggregation device. The authors run a series of experiments and claim that in their experimental pari-mutuel betting markets, information aggregation occurs since the market prices implicitly determined from the market odds are reasonably close to the aggregated information available (the distribution derived from the pooling of all observations).

In fact, their implicit prices are closer to the aggregated information available than the prediction of several models mentioned in their work (Decision Theory Private Information, Competitive Equilibrium Private Information) and the statistics called Average Opinion (the average of individual beliefs before the market opens); only the Best Opinion statistics performs better than implicit market prices, i.e. bettor(s) exist who have

⁵Beliefs are assumed to be completely uninformative, i.e. not correlated with the actual outcome at all. Information, on the other hand, is an informative signal, i.e. correlated with the outcome.

more accurate beliefs before betting than the market does after the period is over.

However, Plott, Wit, and Yang (2003) state that it is not clear why the information is aggregated in their market. The main results of their paper are: First, information was aggregated. This suggests that traders updated their private information based on observed market odds. Second, a model based only on the use of private information (the Decision Theory Private Information model) seems to fit their data best. The authors call this paradoxical.

Kálovcová and Ortmann (2009) replicate their experiment and their results suggest that the paradox seems due to aggregate rather than individual level data analysis. The authors show that market odds (a representation of data aggregated over all participants) are indeed very close to odds that would exist if traders behaved according to the Decision Theory Private Information model. However, the individual level data analysis suggests that, apart from private information, traders extract significant additional information from observing the market odds and behavior of other traders. This process then leads to the aggregation of pieces of information in the hands of several traders.

Roust and Plott (2005) focus their attention on fighting well-documented problems with information aggregation in *pari-mutuel betting markets* (late betting, bubbles). Their experiments show that a special "two-stage" *pari-mutuel* mechanism has the potential to speed up the process through which information is revealed and to reduce deceptive behavior⁶ and incorrect aggregation (informational cascades).

In the first stage, similar to a simple lottery for a fixed prize, participants are endowed with a certain amount of money which they can use to bet on their preferred outcome. This money has no alternative use, and therefore traders have incentive to spend all of their budget and thus reveal all the information they possess. Furthermore, only the aggregated amount of money that is bet on each outcome is observed at the end of the first period, not the individual bets; hence, participants have no incentive for any kind of deceptive behavior.

⁶Strategic behavior based on investing against one's beliefs in order to mislead other traders and therefore affect market odds in a desirable direction.

The second stage is a pari-mutuel betting system with an increasing price of tickets to prevent waiting strategies. This two-stage pari-mutuel mechanism preserves the ability of the betting market to aggregate information and reduces the number and intensity of bubbles. Moreover, after the first stage, participants learn the strength of the signal. If the aggregated information is poor, participants are more likely to rely on their own information which prevents the formation of bubbles and leads to a more reliable aggregation of information.

Regarding the process of aggregation of information, in a light of above mentioned literature, prediction markets seem to be a better choice. Pari-mutuel systems sometimes fail tests of efficiency (see e.g. Bullen 2009). As will be discussed in section 1.5, this inefficiency is most likely related to pari-mutuel betting market rigidity. Money once invested cannot be taken back, and the more money in total that is bet, the lower the effect is of additional bets on the market odds. Therefore, traders who place bets towards the end of the market — it is usually informed traders who wait — have no significant impact on the market odds. On the contrary, in a continuous double auction system or call market, the prices are determined exclusively by the last trades.

The inefficiency of the pari-mutuel betting market is often tackled by the two-stage mechanism (e.g. Roust and Plott 2005). However, the mere fact that the second stage is needed reflects poorly on the ability of pari-mutuel betting markets to aggregate information successfully.

Favorite-longshot bias is closely related to market efficiency and the aggregation of information. This phenomenon is frequently discussed in the literature (see, for example, Sauer 1998 or Thaler and Ziemba 1998), and it describes the situation where shares of favorite events are underpriced, and shares of longshots are overpriced. This gives traders who bet on a longshot a negative expected payoff, and traders who bet on a favorite a positive expected payoff (if there is no house charge).

As Berg and Rietz (2010) suggest, both context and market structure might affect the likelihood of the existence of favorite-longshot bias. The authors do not find favorite-longshot bias in their Iowa Prediction Market data. On the other hand, favorite-longshot

bias is present in, for example, Plott, Wit, and Yang (2003), which we discuss in the second chapter of this dissertation. As Berg and Rietz (2010) suggest, betting markets may be more prone to favorite-longshot bias than prediction markets because of their structure. The two-sided structure of prediction markets (there are both buyers and sellers as opposed to betting markets with buyers only) allows for the easier correction of biased prices. However, this conjecture needs to be looked at more closely and probably experimentally tested in further research.

1.3.2 Insiders

- In a prediction market, insiders are likely to enter the market first so that they can exploit profit opportunities. Prices on an efficient market tend to converge to their true value. Should there be insiders in the market, they would profit more from trading shares early while the prices are off their true values. However, we are not aware of any evidence of this conjecture, and it might be an interesting idea for further experimental research.
- In a betting market, insiders tend to wait and make bets late (Asch, Malkiel, and Quandt 1984). This is so because the payoff on the betting market depends on the amount of money bet in total and, more importantly, on particular event. The higher the betting amount is, the lower the payoff. Therefore, insiders try not to reveal what they know because it could attract more traders. In fact, they might bet on other events first in order to make a last-minute bet on the winner more profitable.
- The situation in a fixed-odds betting market is different. If the odds are set by a market maker, then there is no reason for insiders to postpone their trades towards the end of the market.

Insiders drive the prices towards their "true" values, and therefore an understanding of their behavior on the market is of the utmost importance for a decision maker. Insiders are likely to enter the prediction market earlier compared to a betting market. In our experimental betting market (see Kálovcová and Ortman 2009), we indeed observed a

significant proportion of late betting. In contrast, the experimental internal prediction market that we conducted⁷ was very active in its early stages.

These observations are not surprising: In a betting market, insiders profit the most if the final market odds differ from their true values and hence are reluctant to add their private information to the system. In prediction markets, on the other hand, insiders profit the most from uninformed traders with inaccurate information in the early stages of the market.

1.3.3 Occurrence of Bubbles

- In prediction markets, the probability of creating an informational bubble is lower compared to betting markets. This is implied by the behavior of insiders, who are attracted to come and trade in the early stages of the market, and therefore, the prices reflect the correct information from the early stages of the market.
- In betting markets, insiders profit from postponing their trades till the end of the market and therefore, their information is not reflected in prices till the very last moment, and the likelihood of prices moving in the wrong direction increases.

In more detail, we conjecture that insider behavior makes the betting system more prone to informational bubbles for two reasons.

First, as Ottaviani and Sorensen (2006c) point out, traders who have inside information prefer to bet late not to reveal their private signal. While traders with the best information are rather inactive, other participants use inaccurate market odds based mostly on noisy trading to update their personal information, and informational bubbles are likely to occur.

Second, when informed individuals start participating in the betting process, their power to navigate prices towards the true value is very limited. If many traders participate in the market, the total money bet on all outcomes is large relative to the bets of informed traders. This can harm the accuracy of predictions as insiders cannot drive the wrong

⁷The internal prediction market at CERGE-EI is described in chapter 3.

price to its true value because of their limited impact on the final odds. On the other hand, in a prediction market, the price is determined solely based on the last trade. Even if insiders enter the market late, they can correct the price within one period because the trading history does not affect current prices.

Based on the typical behavior of insiders, we conclude that prediction markets are more likely to gather all the information and less likely to create informational bubbles. However, even in prediction markets not all of the information available is necessarily reflected in market prices.

As Kyle (1985) and Wolfers and Zitzewitz (2005) point out, if the profit of insiders is not high enough (the spread between the ask and bid price is too small, or there are not enough uninformed traders), they do not have an incentive to trade and therefore not all of the information is reflected in the market price. Nevertheless, the same problems are present in betting systems, where insiders do not have incentives to participate in the market if their potential profit is small, i.e. if market odds are very close to their true values. Moreover, apart from the common problems with insufficient motivation for insiders, the betting system also suffers from a high proportion of late trading and consequent market inefficiency. Therefore, from this perspective, the prediction market seems more of an appropriate market mechanism.

1.3.4 Expected Payoff

- In a prediction market, the expected payoff is known to the trader at the moment of trade. It is the difference between the price and a private belief about what the true value is. The payoff from a given trade does not depend on one's own subsequent activities or on the subsequent actions of other traders.
- In a betting market, the payoff will be known only at the end of the whole market. Only when the market is closed, can final odds, and therefore, payoffs be determined. Moreover, the payoff per dollar bet on a winning horse is a *decreasing* function of the total bets placed on that horse.

- In a fixed-odds betting market, the situation is similar to that in a prediction market. The expected payoff is known to a trader at the time of a trade. It is the difference between the odds and a private belief about the likelihood of an underlying event.

1.3.5 Price

- In a continuous double auction prediction market, traders make limit orders to buy or sell given the amount of shares for a given price. If bid price exceeds ask price, trades are matched immediately. Moreover, any trader can accept the most advantageous current order.
- In a call market, the price of a contract is determined such that demand is equal to supply. The price can only be determined if at least some of limit orders to buy are higher than some of the limit orders to sell. Otherwise, the price is not determined, and trades do not occur.
- In a betting market, the price of a ticket is fixed, and the supply is not limited. The final payoff is determined by the demand.

1.3.6 Risk-less Portfolio

- In a prediction market, it is possible to hold a risk-less portfolio. If a trader holds one unit of each contract in the market, he gets 1 currency unit for sure. If the sum of the prices of all contracts is equal to 1, then the trader's net profit is equal to 0. Alternatively, a trader can hold cash as a risk-less investment.
- In a pari-mutuel betting market, it is not possible to hold a risk-less portfolio other than cash. Even if a trader bets one unit of money on each possible outcome, the profit is uncertain. Depending on the actions of other bettors, this strategy can be profitable, but it is also possible that the bettor ends up with a loss.

Consider, for example, a simple pari-mutuel betting market with two bettors and two events: A and B. The first bettor bets one dollar on each outcome, and the second bettor

bets one dollar on A and 3 dollars on B. If A wins, the first bettor gets three dollars back, and his net profit is one dollar. If, however, B wins, then the first player gets 1.5 dollars back and ends up with a net loss of 50 cents. In a pari-mutuel betting market, the final odds relevant for the profit are not known until the market is closed; therefore, traders cannot adjust their bets such as to guarantee sure profit.

The situation is different in the case of a fixed-odds market. In this case, based on the odds, the bettor can bet different amounts of money on different outcomes in such a way that he gets his investment back. Consider a simple fixed-odds betting market with two events, A and B, and odds 3:1 and 1.5:1 respectively. If a bettor bets one dollar on A and two dollars on B, he will get a profit of three dollars irrespective of the winning event. Note, however, that this only holds if the market maker does not take any profit for himself. Otherwise, the value of this risk-less portfolio is negative.

In our opinion, the existence of a risk-less portfolio affects the behavior in the market in two ways. First, some trades are motivated by hedging, and hence, we can expect more active behavior of market participants. Second, we can observe non-zero prices of the shares that correspond to the outcomes that are certain not to occur (as observed in our internal prediction market).

On the betting market, the only motivation to bet on a certain outcome is to gain a profit, and hence, traders bet only on outcomes that are, according to their private beliefs, likely to win⁸.

On the prediction market, traders might want to hold a risk-less portfolio because this portfolio can be exchanged for currency. To hold a risk-less portfolio, a trader has to hold one share of each event and therefore has the motivation to buy the shares of longshot events also. This brings additional liquidity to the market, but at the same time, this type of trading brings additional noise to market prices. This noise, however, affects prices only in the short term. In the long term, the prices of events not likely to occur are pushed to zero.

⁸We exclude attempts at manipulation from this argument because as it has been shown, if possible, manipulation is extremely difficult. See, for example, Camerer 1998 or Rhode and Strumpf 2004.

1.3.7 Trading

- In a continuous double auction prediction market, continuous trading is possible. Traders can buy shares, and they can resell these share later on.
- In a call market, traders can place limit orders continuously, but trades are matched discreetly at predetermined frequency.
- In most betting markets, there is no possibility to resell the ticket once it is bought. The possibility to cancel the bet is very rare, and even then, it is not mentioned in any program information distributed to bettors; therefore, most bettors "were quite surprised that cancelation is possible" Camerer (1998), p.467. Traders can make bets continuously.

Traders can bet against the original event; however, this action will not cancel the original bet entirely. In a pari-mutuel betting market, the final odds relevant for the payoff are not known till the end of the market. So, the bettor cannot know how much he should bet against the original event. He could wait till the very last minute and bet based on current odds, but the odds will still change a little, making it impossible to cancel the bet entirely.

In a fixed-odds market, the “appropriate amount” to be bet against the original event is known, however, as was already mentioned above, the house charge makes this cancelation unprofitable.

1.3.8 Profit for the Market Maker

- In prediction markets, profit for the market maker usually takes the form of a per transaction fee. Although the market maker can choose not to charge transactions, for example in the case of small scale internal markets, to encourage trading.
- In betting markets, the market organizer usually applies a house charge (part of the collected wagers that goes to the market maker as profit). In some cases (usually for experimental or internal betting markets), a negative house charge can be applied to enhance trading and information aggregation (See for example Plott, Wit, and Yang

2003). In either case, the market maker is not at a risk of a loss.

- In a fixed-odds betting market, the odds are set by a market maker who is at risk of losing money should he set the odds incorrectly.

1.3.9 Liquidity

- On small prediction markets, low trading activity can lead to low liquidity, which might be discouraging for traders.
- In betting markets, the liquidity is unlimited, and traders can bet as much money as they want. Sometimes betting takes the form of buying tickets on outcomes. In this case, the supply of tickets is unlimited, and traders can buy as many of them as they want.

1.4 Payment Scheme

In this section, we discuss the effects of two different payment schemes on the accuracy of market predictions. Participants with relevant information have to be motivated to trade actively based on their information. This requires an appropriate payment scheme. Our internal experimental market, discussed in chapter 3, suggests that information is more likely to be aggregated if all traders are paid a "small" profit according to their performance (buying for a low and selling for a high price, or buying the winning events cheaply) compared to the system in which only a few of the most successful traders get the "big" prize.

This observation is consistent with the results in James and Isaac (2000), where the authors discuss the negative influence of tournament prizes on the performance of traders. Their argument is based on the fact that in order to win one of a limited number of prizes, traders have to take risks, and this alternates their behavior often leading to inefficiencies on the market.

The performance of our internal prediction markets also seems to be better if participants are required to (at least partly) invest their own money. However, Luckner

and Weinhardt (2007) present somewhat surprising results. They study the impact of different monetary incentives on prediction accuracy in a prediction market for the FIFA World Cup 2006.

The authors conduct a field experiment with three treatments with the same expected payoff: The subjects of the first group were paid a fixed payment irrespective of their performance. In the second group, three of the most successful individuals were paid according to their ordinal rank. All the other traders in the second group did not receive any payment at all. In the third treatment, subjects' payment linearly depended on the traders' deposit value in the prediction market and was therefore directly influenced by every transaction a trader conducted.

The authors show that performance-related payment schemes do not necessarily increase the prediction accuracy. Due to the risk aversion of traders, the competitive environment in a rank-order tournament leads to the best results in terms of prediction accuracy. Further, even the group with the fixed payment beats the group with the performance-compatible payment.

Another issue related to the payment scheme is whether participants should be required to invest their own money for trading. Using one's own real money seems to have two effects: First, participants are more careful when making trading decisions (with less risky behavior, we observe more diversification in trading) and traders have stronger incentives to participate actively in the market. Second, fewer uninformed traders are attracted to the market if they can lose their own money, and hence, there is more information present in the system.

1.5 The Complexity and Flexibility of Prediction and Betting Markets

We observe that prediction markets are significantly more widespread and being used in forecasting for a great variety of events (political markets, corporate markets, entertainment industry, sports markets), while betting markets prevail mostly in sports betting.

To our best knowledge, the reasons for this difference have not been discussed in the existing literature. In light of their key characteristics, we conjecture that this selection appears mainly due to two reasons: simplicity; and efficiency combined with flexibility.

First, sports betting (mostly organized through the Internet) is intended for the general public, and the betting market environment is somewhat easier to understand. There is only one side of the market (buyers) so no direct interaction between buyers and sellers is needed. Moreover, loosely speaking, given the market odds the only decision to be made is how much money to bet on a favorite outcome. Hence, the betting system appears to be more convenient if we expect mostly unexperienced traders to participate.

On the other hand, trading in a prediction market (typically organized as a continuous double auction or a call market) requires a much better understanding of the market mechanism. Traders need to continuously follow the orders of other traders and decide if they want to accept them or if they want to place their own limit orders. In other words, the prediction market mechanism requires more sophisticated behavior from traders and therefore seems more suitable for more experienced traders, smaller scale markets, or internal markets, where sufficient support can be provided to traders.

On some occasions, prediction markets can be organized in such a way that traders do not submit limit orders. Rather, they submit their subjective probabilities of the likelihood of each event, how confident they are about this likelihood, and the quantity of shares they are willing to buy. Based on this information, limit orders are generated for them. Essentially, the buy order is generated for the most likely event(s), and sell orders are generated for events not likely to occur. The price is determined based on how confident traders are about their belief. One case of this type of market is reported in Polgreen, Nelson, and Neumann (2007). This makes trading in prediction markets much easier; however, it also imposes limitations on the flexibility of traders to place multiple limit orders and to make multiple trades at different prices.

Despite its relative difficulty, the prediction market mechanism seems to be preferred in various forecasting problems. This seems due to its apparent advantages, mainly higher efficiency and flexibility, and this is our second reason for the prevailing use of prediction

markets. As Bullen (2009) points out, in a betting market, the flexibility of market odds decreases significantly as the total money bet in the market increases. Moreover, waiting strategies can be profitable (insiders want to wait to be able to place bets with favorable market odds), and hence, there is no incentive to achieve efficiency before the market closes.

This contrasts with the continuous double auction market or call market in which prices reflect the most recent trade. Moreover, profits can be made at any time, and hence, traders do not wait till the end of the market, and their information is reflected in the prices sooner. Therefore, the aggregation of information is more efficient. All in all, while relatively complicated, the prediction market mechanism seems to be more flexible and more efficient and therefore more likely to be used to forecast a great variety of uncertain outcomes.

The system of trading and the process of how prices are determined are very closely related to the role of insiders and their impact on the final prices. The prediction market is more flexible because traders can resell shares that they bought should they change their beliefs about the probabilities of final outcomes. As a result, the prices of shares of events that are not likely to occur drop towards the end of the market easily because the prices in a prediction market reflect the last trade and therefore the latest beliefs of traders.

The evidence from our short-term internal prediction market supports this hypothesis. In our market designed to predict the number of received applications for a Ph.D. program in economics at CERGE-EI, Prague, the prices of several shares dropped to zero immediately after it was certain that the underlying outcomes will not occur. In a betting system, tickets that are bought cannot be returned nor resold.⁹ This means that even if traders learn during the duration of the market that some outcome is unlikely to occur, the tickets betting on this outcome cannot be returned, and the only way to decrease the odds for a given event is to significantly increase betting on other events.

⁹The possibility to cancel a bet is very rare and usually not known by most traders (Camerer 1998).

1.6 Conclusion

We observe that prediction markets are more widespread and being used in the forecasts of a great variety of events (political markets, corporate markets, entertainment industry, and sports markets), while betting markets prevail mostly in sports betting. To our knowledge, no research provides insight into this configuration. The prediction and betting markets literature provide no justification for the choice of the market type.

We analyzed the key differences between prediction and betting markets, and we conjecture that this selection appears mainly due to two reasons: simplicity and efficiency combined with flexibility. Betting systems are very simple and therefore suitable for large-scale markets with participation of general public and hence a high ratio of unexperienced traders. Prediction markets tend to be more efficient but also more complex, and therefore, appropriate mostly for experienced traders or in small-scale markets, where the necessary support can be provided.

1.7 Appendix

EXAMPLES OF PREDICTION AND BETTING MARKETS¹⁰

- Political events
 - next president, election outcomes (general, senate, governor), next party leader, new EU members, resignation (Iowa political prediction market is analyzed in Berg et al. 2008 and Berg, Nelson, and Rietz 2003)
- Financial bets
 - house prices, interest rates, indices (Dow Jones, FTSE, DAX,...), currencies (exchange rates), macroeconomic indicators (inflation), commodities (gold, oil), tax futures, GDP, CPI, international trade balance (Two inflation prediction markets are reported in Berlemann and Nelson 2005 and Berlemann, Dimitrova, and Nenovsky 2006.)
- Social events
 - Osama bin Laden's capture, US air strike against Iran, Hamas recognition of Israel, Bird flu, terrorist attacks
- Public policies
 - Is it worthy to introduce a new vaccination program? Cost-benefit analysis of policies
- Sporting events (A large body of empirical literature exists that analyzes the data sets from sport betting markets)
- Movie and TV industry
 - movie sales, box office returns, next TV competition winner, Emmy and Grammy awards
- Corporate level indicators
 - sales forecasts, project time schedule, generating new ideas (E.g. Chen and Plott 2002; Ortner 1997; Ortner 1998)
- Weather forecast (hurricane), locating a lost submarine, what drugs will be successful

¹⁰Main sources: www.betfair.com, www.tradesports.com, www.hedgestreet.com

Chapter 2

Understanding the Plott-Wit-Yang Paradox¹

Co-authored with Andreas Ortmann

2.1 Introduction

A (parimutuel) betting market as typically used in horse racing and other sports events, is a system in which all bets are collected and the payoffs are then determined by dividing the total amount of money invested by the amount betted on the winning horse.

In Plott, Wit & Yang (2003), the authors address experimentally two fundamental questions: first, is information aggregated on betting markets? Noting that there is no clear theoretical reason why betting markets should aggregate information at all, the authors report that the implicit prices on their experimental markets are very close to the prices that would exist if all agents pooled their information and made decisions on the basis of the pooled data. This observation suggests that the information in their markets does aggregate. Second, which model explains best how information is aggregated? The theoretical model which seems to fit their data best (the Decision Theory Private

¹An earlier version of this work has been published in *The Journal of Prediction Markets* 3(3). pp. 33-44.

All errors remaining in this text are the responsibility of the authors.

Information, or DTPI, model) does not rely on information aggregation whatsoever. The authors call this paradoxical. We refer to their result below as the Plott-Wit-Yang (PWY) paradox.

We replicate their experiment with minor changes and find, first and like Plott et al., a paradoxical result, which is information is aggregated while the data seem to be explained best by a theoretical model that does not require information aggregation.

We show that market odds are indeed very close to odds that would exist if traders behaved according to the DTPI model. However, our individual level data analysis suggests that, apart from private information, traders extract significant additional information from observing the market odds. The PWY paradox seems due to aggregate rather than individual level data analysis.

We also observe a learning effect: In later rounds traders seem to understand the mechanism of the betting market better and put higher weight on the information contained in the market odds rather than private signals. One plausible explanation is that subjects become increasingly familiar with the laboratory environment.

Finally, we examine the effect of risk-aversion on traders' behavior. We find that the degree of risk aversion does not have any impact on the amount of money traders bet in our experimental market.

In the next section, we discuss the PWY paradox, illustrate our explanation, and formally state our hypothesis. In section 3, we explain design and implementation of our experiment. Results are reported in section 4, and we conclude in section 5.

2.2 The PWY Paradox

The paradox consists of two results that contradict each other: first, information is aggregated on the market, i.e. traders are involved in some sort of strategic behavior. Second, if we want to simulate the behavior of traders, our best bet is to use the DTPI model which is based on the use of private signals only. Our explanation of the PWY paradox is based on a detailed analysis of the second result. We show that, while aggregate

level data might suggest that traders follow the DTPI model, individual data analysis might lead to a different conclusion because two different trading behaviors can lead to the same aggregate results. Our argument can be illustrated by the following example:

Example: Suppose that there are only two traders on the market (Trader 1 and Trader 2) with the same budget and only two ex-ante equally likely events A and B that traders can bet on. Further suppose that based on their private signal, Trader 1 thinks that A is the winning event, and Trader 2 thinks that B is more likely to win. If both traders behave according to the DTPI model, then every trader invests all the money into the more likely event and the resulting market odds are 2:1 for both events A and B. Alternatively, traders can behave strategically and by observing the other trader's actions, they learn about each other's information. Consequently, they both invest half of the budget into each event. Again, the resulting market odds are 2:1 for each event.

When searching for an underlying model, Plott, Wit, and Yang (2003) look at the aggregate level data and conclude correctly that the prediction of the DTPI model fits market odds the best. In terms of the example above, Plott et al. observe market odds 2:1 for each event and claim that the prediction of the DTPI model is also 2:1 for each event, and hence, traders follow this model. Our evidence suggests that it is not necessarily true.

Hypothesis: *Traders do take into account information contained in their private signals and information contained in market odds.* In other words, traders observe behavior of others and based on market odds, they update private beliefs. Through this process information is aggregated and translated into market odds.

2.3 Our Experimental Betting Market

Because the original data are lost, we replicate the Plott, Wit, and Yang (2003) design. We change minor implementation details all designed to allow us to analyze the impact of risk-aversion.²

²The experiment was programmed in z-Tree Fischbacher (2007). Instructions for this experiment and the data can be found at: <http://home.cerge-ei.cz/kalovcova/research.html>

2.3.1 Design

The design of our betting market follows the one in Plott, Wit, and Yang (2003). Subjects bet on six events labeled A, B, C, D, E, and F which are equally likely *ex ante*. In each round, one of the letters is drawn at random from an urn, recorded, and then placed back into the urn. In other words, the draw of an event is independent across rounds, and the history of draws holds no implications of what future draws might be. Which of the events wins is announced after the end of each round.

After the winning event is chosen, each individual is privately given a noisy signal (or "clue") about the winning event. The clues are determined independently for each individual by the following procedure. Once the winning event is determined, a new urn is created with five letters of the winning event and two letters from each of the other events. The participant is informed of the outcome of three random draws with replacement. The information distributed across all participants in a session is more than that of any one individual. However, this information is not sufficient to determine the winner with absolute certainty.

2.3.2 Implementation

Employing 109 undergraduate students, we conducted our experiment in four sessions in February 2008 and an additional five sessions in March 2009 each of which consisted of one trial round (which did not affect the earnings and was intended to make subjects familiar with the software) and then continued with 16 regular rounds.

Time, in seconds, was displayed on each computer screen. The duration of each round was 120-300 seconds - the time of duration was chosen randomly and independently for each round and was unknown. At the end of the experiment, four rounds were randomly chosen, and subjects were paid based on their performance in the paying periods.

The price of each event ticket was 1 ECU (experimental currency unit), once a ticket was bought it could be neither returned nor resold. At the beginning of each round subjects were endowed with 300 ECU, which they were free to spend or to keep. The

part of the endowment not spent declined in value as subjects were allowed to keep only three-quarters of it. After subjects spent their entire endowment, they could get a loan of 600 ECU, which had to be paid back after the end of each round.

The payoff for each round was determined in the following way:

$$\begin{aligned} \text{Payoff} &= 0.75 \times \text{money on hand (part of the endowment or loan not spent)} \\ &\quad + \text{profit} \\ &\quad - \text{loan payback (if the loan was taken),} \end{aligned}$$

where

$$\text{profit} = \frac{\text{Total ECU from all ticket sales}}{\text{Total number of winning tickets sold}} \times \text{Number of winning tickets held.}$$

The implementation of our experimental betting market differs from that in Plott, Wit, and Yang (2003) in four respects. All four changes served the additional purpose of creating a betting market in which we could observe the effect of risk aversion. That risk aversion might have an impact is strongly suggested by the literature. For a recent and comprehensive review see Harrison and Rutström (2008).

We believe, and the evidence below suggests, that these differences in implementation do not affect the participants' behavior to the extent that is relevant for an examination of the PWY paradox. In the section below, we discuss the implementation changes. In the results section, we discuss briefly the effects of the risk aversion and the specifics of the risk-aversion instrument we used. We focus, however, mostly on results directly connected to the PWY paradox.

1. Risk-aversion. To measure the level of risk aversion, we administered the assessment instrument proposed in Holt and Laury (2002) and now widely used for that purpose. Participants were financially incentivized for this part of the experiment. To control for the order effect, we administrated the risk aversion measure prior to the betting market part of the experiment in the first four sessions and after the betting market part in the next five sessions of our experiment.

2. *House bonus.* In Plott, Wit, and Yang (2003) a house bonus is used. A house bonus is the money added to the total amount of money invested by all the subjects. The expected payoff from the investment is thus strictly positive and gives risk-averse subjects better incentives to invest: The house bonus makes investment more profitable and the more subjects invest the more information can be aggregated. The house bonus seems responsible for the successful information aggregation on the betting markets in Plott, Wit, and Yang (2003).

However, this mechanism calls for investing in the market as much as possible. Hence, the traders who fully understand this mechanism invest all the money they have irrespective of their attitude towards risk. Only extremely risk-averse subjects would do otherwise. In real betting markets, a house charge is used instead of a house bonus. The expected payoff from the investment is then slightly negative, and risk-aversion is likely to play a significant role. In our experimental betting market, neither house bonus nor house charge is used.

3. *Endowment depreciation.* Without a house bonus, traders are less motivated to invest in the market. Pilot experiments that we conducted confirmed this hypothesis: Subjects spend 78% of all the money at their disposal in markets with a house bonus, whereas they spend only 64% in markets without a house bonus. To enhance the process of information aggregation, we wanted to make sure that traders would spend a major part of their endowment.

Plott, Wit, and Yang (2003) use an experimental design in which the part of the endowment that is not spent is lost. This makes all subjects spend the entire endowment. In our betting market, the part of the endowment that is not spent declines in value, and subjects are allowed to keep only three-quarters of it. This design creates strong incentives for subjects to spend a major part of the endowment and thus allows for information aggregation. At the same time, extremely risk-averse participants are allowed to keep all of the endowment and earn a small but sure profit. Hence, we enhance the process of aggregating the information while keeping risk-aversion to play as a significant aspect.

4. *Paying periods.* After completing the experiment, we randomly selected four periods for which subjects were paid and this was ex-ante known to all participants. (In Plott, Wit, and Yang (2003) subjects were paid in all rounds.)

We implemented this payment mode to prevent subjects doing nothing and only shortly before the market is closed investing all their money into the event with the lowest odds, i.e. the most likely event. Most of the time the true event is identified successfully, and hence, the sniping strategy that we observed in pilot markets leads to a large long-run profit (negative profit in a few periods is offset by a large positive profit in most of the periods).

However, extensive waiting worsens information aggregation because subjects who wait keep their private information away from the market. With our payment mode, waiting and investing everything in the most likely event becomes less attractive because those periods where the profit is negative could be chosen to be paying periods.

We realize that based on the expected utility theory, this payment scheme should not alter traders' behavior at all. If the expected payoff from sniping is positive in one period, then it is positive in all periods and vice versa. However, based on our pilot experiments where we observed a significant amount of sniping and based on ex post discussions with our pilot market subjects, we decide to employ this payment scheme. In subsequent experiments, the sniping significantly decreased. The reason might be connected to the loss aversion of our subjects.

However, we did not analyze this issue more closely because this was not the area of our research. The bottom line is that we implemented our market in such a way as to give information aggregation the best chance. Our payment scheme either helped limit undesirable sniping or did not affect the timing of bets at all, and late betting observed in pilot markets was just unusual behavior.

2.4 Results

We start this section with two results that constitute the Plott-Wit-Yang paradox. Then we follow with the third - key result - that supports our hypothesis, and we finish with three supplementary results concerning learning effect, market efficiency, and the effect of risk-aversion.

Result 1: Information is aggregated. Similar to Plott, Wit, and Yang (2003), we find evidence in favor of information aggregation. The results are provided in Table 1 below in the form of the Würtz³ measure of the distance of model predictions from AIA (Aggregated Information Available, i.e. posterior probabilities given the pooled signal of all traders). The Würtz measure is computed for aggregate data. For example, the Würtz measure of distance between DTPI and AIA is determined in the following way: First, we compute what market odds would be if all traders behaved according to the DTPI model and determine the corresponding probability distribution p_i . Then we take the probability distribution given by AIA, q_i , and use the formula in Footnote 2 to compute their distance.

In Table 1 below we follow the notation in Plott, Wit, and Yang (2003):

- Decision Theory Private Information Model (DTPI) - a model where traders base their decisions exclusively on their own private information and bet all their money on the most likely event.
- Competitive Equilibrium Private Information Model (CEPI) - a model where traders take market odds as constants and maximize their expected profit with respect to their private information.⁴
- Average Opinion statistics (AO) - the average of individual beliefs before the market opens.

³If the discrete distributions are described by their probability density functions $\{p_i\}_{i=1\dots K}$ and $\{q_i\}_{i=1\dots K}$ respectively, then the measure proposed by Wurtz (1997) can be written as $W(p, q) = 0.5 \sum_{i=1}^K |p_i - q_i|$.

⁴We use the method described in Eisenberg and Gale (1959) and *Mathematica* to compute equilibrium odds.

- Best Opinion statistics (BO) - the most accurate belief among traders' beliefs before betting.
- Implicit prices (IP) - market prices implicitly determined by the market odds.

Table 1: Average Würtz measure of distance from AIA:

all periods:				
Best Opinion	IP	DTPI	Average Opinion	CEPI
0.380 (0.163)	0.495 (0.193)	0.515 (0.102)	0.634 (0.099)	0.663 (0.101)
last 8 periods:				
Best Opinion	IP	DTPI	Average Opinion	CEPI
0.427 (0.187)	0.489 (0.191)	0.511 (0.111)	0.627 (0.105)	0.657 (0.109)

The results in Table 1 show that the distribution of probabilities based on IP is closer to the distribution given by AIA than the prediction of any other model except BO. For example, in the first row of Table 1, the Würtz measure of the distance between AIA and IP is 0.495, which is lower than the Würtz measure of the distance between AIA and any other model except BO. This means that apart from BO, IP is closer to the AIA than the prediction of any other model.

We also observe some indication that information aggregation can improve over time. At the beginning of the experiment, the information aggregation is weaker, which is probably caused by the inexperience of participants (the average Würtz measure of the distance between AIA and IP across all periods is 0.495). As the experiment continues, participants understand the mechanism better, behave more strategically, and try to update their own signal based on what happens on the market. Hence in later periods, the information aggregation is more obvious (the average Würtz measure of the distance between AIA and IP across the last 8 periods is 0.489). Due to the low number of observations, the difference in Würtz measure during the first and last 8 periods is not statistically significant.

Result 2: DTPI model best fits the data from betting markets.

Table 2: The average Würtz measure of distance of model predictions from IP:

all periods:				
DTPI	Average Opinion	CEPI	Best Opinion	AIA
0.261 (0.124)	0.306 (0.193)	0.330 (0.134)	0.324 (0.161)	0.495 (0.134)
last 8 periods:				
DTPI	Average Opinion	CEPI	Best Opinion	AIA
0.269 (0.126)	0.330 (0.123)	0.355 (0.120)	0.340 (0.169)	0.489 (0.191)

The average Würtz measure of the distance between IP and DTPI across all periods (0.261) and across the last 8 periods (0.269), is lower than the average Würtz measure of the distance between IP and any other model. This means that the DTPI model best fits the experimental data. Results 1 and 2 are in line with the results in Plott, Wit, and Yang (2003) and constitute the PWY paradox. In the following section, we provide an explanation for this paradox.

Result 3: Our PWY paradox explanation is supported by the data.

First, we find that traders invest on average one-third of their overall investment into events that they should ignore according to DTPI model. Second, we compare the observed individual distribution of bets to the distribution of bets implied by private signals (the DTPI model) and the distribution of bets implied by market odds (bets are in proportion to their probabilities implied by market odds). For this comparison, we use again the Würtz measure.

Our hypothesis is that traders take into account information contained in their private signals and in market odds, which implies that the Würtz measure (Würtz criterion, WC) of distance between observed individual behavior and the private signal is approximately the same or larger than the WC of distance between observed individual behavior and behavior induced by the market odds ($WC[\text{Behavior-Signal}] \geq WC[\text{Behavior-Odds}]$. Note that the smaller WC, the shorter is the distance between the two distributions).

We run a t-test on our data and find extensive support for this hypothesis. We can reject the null-hypothesis that the Würtz measure between observed behavior and market odds is the same as the Würtz measure between observed behavior and a private signal ($WC[Behavior-Odds]=WC[Behavior-Signal]$) in favor of the alternative hypothesis that the Würtz measure between observed behavior and market odds is lower than the Würtz measure between observed behavior and a private signal ($WC[Behavior-Odds]<WC[Behavior-Signal]$) at any reasonable level of significance (p-value is 0.00). In other words, we find support for the fact that traders rely on the signal contained in market odds more than they rely on their private information.⁵

Moreover for every trader, we analyze separately periods with a strong signal (at least two out of three draws are the same; i.e. the probability of the most likely event is 50% or 75%) and a weak signal (all three draws are different; i.e. three most likely events are equally likely with probability to occur equal to 24% each). We find that traders follow market odds more closely than their private signal irrespective of the quality of their private signal. We can reject the null-hypothesis, $WC[Behavior-Odds]=WC[Behavior-Signal]$, in favor of the alternative, $WC[Behavior-Odds]<WC[Behavior-Signal]$, at any reasonable level of significance (p-value is 0.00 in both cases).

To provide an additional insight into the data, we analyze the group of rounds in which the private signal is in line with market odds and the group of rounds in which the private signal and market odds differ. We find that

- in the first group, $WC[Behavior-Odds]=WC[Behavior-Signal]$ (p=0.10 with a two-sided alternative hypothesis; p=0.05 with a one-sided alternative hypothesis).
- in the second group, $WC[Behavior-Odds]<WC[Behavior-Signal]$ (p=0.00).

Therefore, we conclude that traders form a weighted average of their private signal and market odds with approximately equal weights when their signal is consistent with market odds. However, traders trust their private signal significantly less if it contradicts the market odds.

⁵We also ran a non-parametric Wilcoxon rank-sum test. The results are qualitatively the same.

Result 4: Traders' behavior is increasingly influenced by public signals.

We analyze the first and the last eight periods of our experiment separately, and we find support for a learning effect: We find that traders follow market odds more than private signals in the first eight periods, and they rely on market odds even significantly more during the last eight. In particular, traders follow private signals slightly less in latter periods (however, this result is not statistically significant), and secondly, traders follow market odds significantly more in latter periods (p-value is 0.00).

These results suggest that after the traders understand the mechanism of betting markets better and learn that the market works well identifying the winning event, traders shift weight towards the public information in the form of market odds.

Result 5: Market is efficient. The betting experiment exhibits weak statistical efficiency.

Table 3: Winning probabilities assigned by the betting market and actual frequencies of winning.

Market Rank by IP	Average IP	Frequency of Winning	Standard Error of Frequency of Winning	t-statistics
1 st	0.517	0.660	0.150	-1.060
2 nd	0.191	0.132	0.077	0.522
3 rd	0.108	0.125	0.044	-0.205
4 th	0.078	0.069	0.034	0.250
5 th	0.060	0	0.030	2.014
6 th	0.046	0.014	0.026	1.220

In Table 3, markets are ranked according to the average implicit price (IP) for all sessions. The average IP of the 1st market is 0.517. Actual relative winning frequency of the 1st market is 0.660. We cannot reject the null-hypothesis, the two distributions (column 2 and column 3) are the same. As a result, we cannot reject the weak statistical efficiency of this betting market with the exception of the 5th market for which the implicit price is significantly larger than the actual frequency of winning. Hence, the efficiency of the market is not so profound. We also observe a favorite long-shot bias in

our markets: the market probability for favorites is understated (0.517 with an actual winning frequency of 0.660) and the probability for long-shots is overstated (0.046 with an actual winning frequency of 0.014). However, this result is not statistically significant.

Result 6: Risk Aversion Does not Affect the Level of Investment.

Out of 109 participants in our experiments, there were 24 participants for whom the level of risk-aversion could not be measured, and they were omitted from further analysis.⁶ We divide the remaining 88 participants (35 from the first and 50 from the second round of experiments) into two groups - 51 more (15 from the first and 36 from the second round of experiments) and 34 less (20 from the first and 14 from the second round of experiments) risk averse participants.

The null-hypothesis, that less and more risk-averse participant spend on average the same amount of money, cannot be rejected at any level of significance (p-value is 0.9) in favor of the alternative hypothesis, less risk-averse traders spend more money. We do not observe any significant difference in risk-aversion distribution (the p value in the Wilcoxon-Mann-Whitney rank sum test is 0.43).

If we analyze the 2008 data, where risk-aversion measure comes first, separately, we find that less risk-averse participants spend on average 20% more than more risk-averse individuals. The null-hypothesis, less and more risk-averse participant spend on average the same amount of money, can be rejected at the 10% level of significance (p-value is 0.9) in favor of the alternative hypothesis, less risk-averse traders spend more money. We can say that more risk-averse individuals will participate less, and hence, their private information will have less of an impact on implied prices, with a resulting loss in efficiency.

If the risk-aversion measure comes second (2009 experiment), participants invest on average the same amount irrespective of their risk aversion. We tested for the order effect of the risk aversion assessment instrument, and we did not find any differences in the participants' distribution of risk aversion among the first series and the second series of sessions.

⁶These individuals made multiple switches between the safe and risky lottery.

2.5 Conclusion

We replicated the experimental betting market in Plott, Wit, and Yang (2003). Our data confirm the Plott et al. findings on their level of analysis. Specifically, our analysis showed that aggregate data suggest that traders follow the DTPI model. Individually, traders take into account information and the behavior of other traders in the form of market odds, though. Based on this finding, we explained the PWY paradox. Furthermore, we found a learning effect on our betting market. In later rounds, traders put less weight on their private signal and rely more on the signal contained in market odds. Finally, we do not find any effect of a degree of risk aversion on traders behavior.

Chapter 3

A Prediction Market for Applications¹

3.1 Introduction

Prediction markets (also known as information markets, decision markets, or idea futures) are speculative markets created for the purpose of making predictions. Participants trade assets whose final cash value is tied to a particular event (e.g., will the next US president be a Republican) or parameter (e.g., total sales next quarter). Current market prices can then be interpreted as predictions of the probability of an event or the expected value of a parameter.

Evidence cited below suggests that these markets are at least as accurate as other institutions predicting the same events with a similar pool of participants. The reasons are manifold: First, prediction markets are often able to aggregate information that is dispersed (Chen and Plott 2002 and Forsythe, Neslon, Neumann, and Wright 1992) and provide accurate forecasts in spite of individual biases and errors (Forythe, Rietz, and Ross 1999); second, anonymous trading makes participants more likely to reveal what they really know (Berg, Nelson, and Rietz 2003); and third, well-constructed prediction markets are difficult to manipulate (Rhode and Strumpf 2004).

¹The author would like to thank Andreas Ortmann and Peter Katuščák for comments and Jan Mysliveček for substantial help with the implementation of this project. Support through GDN grant No. RRC VIII-89 is gratefully appreciated. All errors remaining in this text are the responsibility of the author.

Economists have started to investigate the potential use of prediction markets in companies' decision making. Indeed, a small but growing number of companies have been creating their own prediction markets over the past decade. Robert Charette mentions several examples in his article:² French Telecom Group, for example, created a prediction market to analyze technological trends; Google has created an internal prediction market to forecast product launch dates, new office openings, etc.

Similarly, Hewlett-Packard created its own internal market system to help predict critical business issues such as the quarterly sales forecast or the price of certain products in one, three, or six months. HP has found that its internal market predictions are often more accurate than the company's "official" forecasts (for instance, six out of eight times, the market was better at predicting computer sales).

Many other companies are experimenting with or creating different types of internal prediction markets: Microsoft (for predicting product shipping dates); GE (for discovering new ideas); Eli Lilly (for discovering new drug candidates); and Siemens (for improving the accuracy of product developments).

Theoretical research has limited possibilities in analyzing the use of prediction markets within corporate businesses or institutions. It has been demonstrated through laboratory experiments that properly designed markets can aggregate information. However, it is also clear that this ability is closely related to market design and implementation. Therefore, field experiments present a very useful tool for analyzing and understanding better internal prediction markets. In this paper, we report the results of our prediction market implementation.

We conducted an experimental prediction market at CERGE-EI, Prague, to forecast the number of applications received for the 2009/2010 Ph.D. program. Our goal was to examine if this specific prediction market can work in a sense that it attracts enough informed traders and reflects their collective beliefs in prices. We believed that a little was known about the plausible number of applications by any single individual, but the

²An Internal Futures Market, *BI Review Magazine*, February 28, 2007. <http://www.information-management.com/bissues/20070301/2600311-1.html>. Accessed on November 6, 2009.

aggregation of pieces of information might be considerable. Therefore we decided to implement a prediction market within our institution to examine the capacity of prediction markets to aggregate information in this very particular setting.

To the best of our knowledge, our experimental market is the first within an educational institution, or for that matter, within a nonprofit organization. We do not overlook the crucial pioneering role of the Iowa prediction market analyzed in Berg, Forsythe, Nelson, and Rietz 2008. However, we feel that our internal prediction market is unique in the sense that the object of our forecast was an internal issue with the number of Ph.D. program applications, unlike in the Iowa market, in that the issue of interest was national presidential elections (or other similar markets conducted for education purposes at various universities, where the outcomes of various sports events are forecasted). In these markets, the information that is being aggregated might already be public knowledge (through news, surveys, or national polls). In our internal prediction market, we try to aggregate bits of private information held by individual traders.

In our market, traders were students and faculty members who were interested in participating and who might have felt that they have insider information about the interest in this program. For example, some students participated in recruiting activities; faculty members also helped in recruiting and knew about extensive financial support for new students, etc. The hope was to establish the prediction market as a leading indicator of the recruitment process that potentially could affect activities in the weeks leading up to the deadline and the ensuing selection process.

3.2 Literature Review

For internal markets to work well, a number of requirements must be met. Internal markets are usually small scale markets and hence the problem with insufficient liquidity has to be overcome. Market participants need to have sufficient information about the matter of interest. Manipulation can also be a problem in a small-scale market. Another open question is that of the optimal role of the market maker — should we, for example, create

an artificial trader to get rid of arbitrage? Another issue is the choice of an appropriate benchmark to measure the performance of our predictions. Unlike in laboratory experimental environments, the information held by all individuals is not known, and therefore, information aggregation cannot be measured.

Little is known about the proper design and implementation of such a market in organizations in general and nonprofits in particular. Hence, the main purpose of the project is to run a trial market and explore its potential. In the following section, we summarize previous experience with the implementation of similar internal prediction markets. Based on this experience, we designed our internal prediction market.

In our review, we concentrate on the literature of similar markets to that which we conducted. In particular, we analyze either small scale or internal prediction markets. We do not ignore large scale markets with widely popular events being predicted (Berg et al. 2008, Cahlík et al. 2005); however, for our purpose, the discussion of the design and results of small and internal prediction markets is more fruitful.

Chen and Plott (2002) conducted an internal sales forecasting prediction market in Hewlett-Packard. Their market was open during lunch and in the evening every day over one week. The market was organized as a double auction. There were 7-26 active participants who were trading with 10 contingent contracts. The authors note that they chose a complete set of state contingent contracts because experiments showed that single compound securities can have difficulty with information aggregation. Participants of this internal market received "small amounts" of cash and were also allowed to invest limited amounts of their own money. The authors show that the prediction market performed better than the traditional methods employed inside Hewlett-Packard.

Ortner (1997) reports on the results from a prediction market created in Siemens. This market was created to estimate the expected delay of certain projects. This double auction market was fully computerized and participants were trading two contracts (project will finish in time, project will not finish in time). Out of 63 participants, 50 traded actively on a regular basis during 44 weeks of duration of the market. Participants used real money on this market. They received 200 ATS after they invested 100 ATS of

their own. Prices on the market converged to the stable equilibrium in about a month. The authors claim that the results show a satisfying forecasting potential.

Berlemann and Nelson (2005) conducted an inflation prediction market in Germany, which was organized as a winner-takes-all market. Using real money, 31-44 traders participated in this market, which was open for 3-5 months and traded eight contracts. According to the authors, the evidence from their pilot market is insufficient to prove that forecasts generated by experimental forecasting markets are of sufficiently high quality to be used in practice. Therefore, they concentrate on how fast market prices reflect the new information, and they find their markets efficient.

Berlemann, Dimitrova, and Nenovsky (2006) summarize the results of predicting inflation in Bulgaria in a similar market. The long-term market was created in combination with short-term markets in order to increase motivation for the traders. Markets with a 3-month and 1-year horizon were open simultaneously for a period of two weeks. There were 25 expert participants who traded 8 contracts. No real money was used. Traders received experimental money, and the three most successful traders were awarded prizes. Their markets were not very successful probably due to an insufficient number of traders, a problematic and insufficient incentive system, a problematic market set-up, and the sorry state of data availability in Bulgaria, among others.

To sum up, previous internal prediction markets were conducted with mixed success. Although market design differs in some details, the basic features of all the markets are the same. We follow this basic design and implementation.

3.3 Our Experimental Prediction Market

Our markets were designed to predict the number of valid applications received for the 2009/10 Ph.D. program at CERGE-EI. The deadline for submitting applications was February 28, 2009. We created two prediction markets. The first, a long-term prediction market, was open for 35 days from January 23 to February 26, 2009 and was designed to predict the number of all applications. The second, a short-term market, was open for 5

days from February 23 to February 27, 2009 and was designed to forecast the number of applications from Czech and Slovak applicants only.

3.3.1 Design

Our markets were organized as winner-takes-all³ call markets with one or two calls per day. In other words, all possible outcomes were partitioned into a small number of subsets. Each sub-set was then related to a security. After the final outcome was known, the winning security, the one which contained the final outcome, was announced. This winning security paid 1 Virtual Dollar per share after the experiment was over. All other securities paid nothing.

The basic structure of our prediction market is very similar to markets cited in Chen and Plott (2002); Ortner (1997); Berlemann and Nelson (2005); and Berlemann, Dimitrova, and Nenovsky (2006). One significant difference lies in the fact that we used the call market rather than the continuous double auction employed by the above mentioned authors.

Our way of clearing the market once (twice) a day instead of continuous trading gave us the following advantage: The markets were open all day, so traders had enough time to place their orders, and they were not limited to using the computer only during a short period of time. At the same time, traders were not discouraged by the lack of activity on the market, which could be an issue should we employ the standard double auction market with continuous clearing of matching orders.

Chen and Plott (2002) solved the potential low level of activity by keeping the market open only during the lunch break. This can work very well inside a corporation where employees take lunch breaks at the same time; however, we think that this time schedule would be problematic for students and faculty members with highly irregular schedules. Hence, we decided to keep markets open all day to increase the number of active traders.

³Note that winner-takes-all market refers to the type of assets traded not to the system of pay-offs/prizes for traders.

Participation

The selection of participants is of utmost importance. We did not want to miss any individual with a relevant piece of information. On the other hand, the participation of many individuals with no useful knowledge is not desirable either.

In the internal markets mentioned in the literature section, participants were either students (Berlemann and Nelson 2005), experts (Berlemann, Dimitrova, and Nenovsky 2006), or employees (Chen and Plott 2002, Ortner 1997). Markets with students and employees performed well, while the Bulgarian inflation market (Berlemann, Dimitrova, and Nenovsky 2006) with expert traders failed to provide accurate predictions. The reason seems to be mainly a problematic and insufficient incentive system and problematic market set-up rather than the choice of participants. This experience suggests that traders do not have to be experts or have previous experience as long as they possess some relevant information and create sufficient liquidity on the market.

In our markets, participation was voluntary. It was not limited in any way. The opening of the markets was announced publicly, and all students and faculty members were invited to participate. We did not impose any restrictions on participation; therefore, our participants self-selected themselves for the experiment.

We are aware of the potential problem of including too many uninformed traders, who drive the market prices away from their true values. However, if we had used some selection process for the participation, we might have had a problem with a low number of participants and consequently with low liquidity on the market. We consider insufficient liquidity to be a more serious issue that would hinder the information aggregation process significantly.

An effort was made to make participation anonymous; however, since most of our subjects know each other very well, we would expect some interactions among them during the experiments. Participants were able to talk to each other during the experiment; however, the trades were made anonymously. The traders were able to participate in one of the markets or both; we did not impose any restrictions on participation.

When the long-term market was open, there was no indication to traders about the short-term market being opened later. The interest in the short-term market was lower compared to the long-term market probably because of the fact that traders participating in the short-term market were asked to invest their own money. All 11 traders from the short-term market were already participating in the long-term market.

Currency and Endowment

In the first, the long-term prediction market, participants were endowed and traded with experimental money, and after the market was closed, the three most successful traders received monetary prizes (380 EUR, 150 EUR, and 75 EUR). This tournament payment scheme was employed by Berlemann, Dimitrova, and Nenovsky (2006).

In the second, the short-term prediction market, participants had to invest 12 EUR of their own money, and this money was used for trading. The net earnings of each participant was equal to the final value of portfolio minus the 12 EUR investment. The net earnings were then multiplied by a factor of 5 to encourage trading. A similar payment scheme with real money and some form of subsidy (usually multiplying a trader's initial investment by some factor) from a market maker was used in markets, which are discussed in Chen and Plott (2002); Ortner (1997); and Berlemann and Nelson (2005).

We started with the design where no real money was used because we were concerned about the participation rate if traders have to invest their own money. Our concern was confirmed by a much lower participation in the second market requiring an investment of the participants' own money. We might argue that the difference in the participation rate, in particular the much higher participation in our long-term prediction market, was caused by high potential prizes in our long-term prediction market. However, the fact that we multiplied net earnings in our short-term prediction market by a factor of 5 made investing in this market very profitable, and a positive profit is much more likely compared to the limited number of prizes in the long-term market.

A prediction market with no real own money invested and with tournament prizes has two potential negative impacts. First, while we did have a sufficient number of

participants in our long-term prediction market, using virtual money attracts many uninformed traders because there is nothing to lose, and hence, the forecast accuracy is not overwhelming.

Secondly, the design of the first market allows only three of the most successful traders to win a prize. This tournament design suggests to play an all-or-nothing strategy, where a trader chooses the most likely event according to their information and invests all the money into buying the corresponding shares. Basically, this is the only possible winning strategy because traders with high but not the top earnings get no prize at all.

This kind of behavior was experimentally examined for example in James and Isaac (2000) who show that tournament incentives lead to risky individual behavior. The authors argue that in order to be better than average, traders have to do something different — they have to take risks. Nalebuff and Stiglitz (1983) draw similar conclusions and claim that tournaments cause traders, who would otherwise behave in a risk-neutral or even risk-averse way, to behave as if they were risk-loving. The tournament payment scheme was implemented for example in the Bulgarian inflation prediction market discussed in Berlemann, Dimitrova, and Nenovsky (2006). We suggest that this might be one of the reasons why this market was not so successful.

The risky behavior which is a result of the tournament payment scheme might be undesirable for a market maker who is interested in aggregated information because traders choose their preferable event a priori and do not update their beliefs nor do they try to gather additional information from the behavior of other traders. In other words, we can learn the prior distribution of traders' beliefs, but the beliefs of an insider and a random trader have the same weight. Usually, through the process of learning and aggregating information, insiders earn money based on their high quality information, and uninformed traders change their beliefs in the right direction. These two reasons led us to change the incentivization scheme of the second market.

3.3.2 Implementation

For both experiments, the market mechanism was web based. We developed and implemented web-based, call market software, which ran on our internal server at CERGE-EI. Traders had access to the entire market through a single web page. For an illustration, see the screenshots in the Appendix.

When the participants logged in to the long-term market, they received a cash balance of virtual \$200 and 200 market portfolios (i.e. 200 shares of each event). The full set of instructions was available to subjects the entire time and they were encouraged to contact experimenter if they encountered difficulties.

Participants had access to all the information about the current round, their current share holdings, the amount of cash and the number of portfolios that they owned, and they were able to place orders to buy/sell individual shares/market portfolios. They also had access to detailed information about previous rounds (orders that they placed, trades that were realized, past market prices, etc.). For details on how this information was provided to the traders, see the Appendix of this chapter.

Market clearing

In each trading period, all orders to buy and sell shares for a particular event were matched up to determine the trades and associated trading prices for shares. All orders from the traders' List of Current Trading Orders were taken into account. Our long-term prediction market was cleared daily at 3 a.m. The short-term prediction market was cleared twice a day at 3 a.m. and 3 p.m.

Details on the process of determining the market clearing price (in the case there was excess demand or supply and in case of ties) can be found in the Appendix.

Events and Participation

The long-term market was open for 35 days or trading periods. There were 57 participants who registered on the market, and they were trading with the following shares:

Event	Each share pays \$1 if the number of applications is
Event A	≤ 80
Event B	81-90
Event C	91-100
Event D	101-110
Event E	111-120
Event F	121-130
Event G	131-140
Event H	141-150
Event I	>150

Figure 1. Definition of events on the long-term prediction market.

The short-term market: was open for 5 days or 10 trading periods. Eleven participants were registered on the market, and they were trading with the following shares:

Event	Each share pays \$1 if the number of applications is
Event A	≤ 15
Event B	16-17
Event C	18-19
Event D	20-21
Event E	22-23
Event F	24-25
Event G	26-27
Event H	≥ 28

Figure 2. Definition of events on the short-term prediction market.

3.4 Results

The long-term market was open on January 23, and it was closed on February 26. However, during the last trading day, there was almost no activity on the market, and with the exception of three securities (namely F, G, and H) no trades occurred. Therefore in our analysis, we only work with the data until February 25, which was the 34th round of the long-term market. The short-term market was open on February 23 and closed on February 27. This market was open for 5 days or 10 periods (the market was cleared twice a day).

In the table below, we show the actual number of applications during the last few

days of our markets (and corresponding events). The end of each market is depicted by a double horizontal line. We also provide the final number of applications received as of March 2, which was four days after our markets were closed and which was relevant for determining the winning events and consequently the payoffs.

We skip weekends in our table because applications, which were the subject of our prediction markets, were delivered by regular mail only during business days: Monday-Friday.

	Long-term Market	Short-term Market
Date	Total Number of Applications (Cumulative)	Number of Applications for CR+SR (Cumulative)
Feb 19	87 (Event B, Period 28)	n.a.
Feb 20	96 (Event C, Period 29)	n.a.
Feb 23	113 (Event E, Period 32)	11 [4+7] (Event A, Period 1,2)
Feb 24	120 (Event E, Period 33)	12 [4+8] (Event A, Period 3,4)
Feb 25	125 (Event F, Period 34)	14 [6+8] (Event A, Period 5,6)
Feb 26	132 (Event G, Period 35)	16 [7+9] (Event B, Period 7,8)
Feb 27	136 (Event G)	19 [7+12] (Event C, Period 9,10)
Mar 2 (Final)	140 (Event G)	21 [7+14] (Event D)

Figure 3. The actual number of applications received during the last days before the long-term and short-term prediction markets were closed.

Activity on our Markets

We had 40 active⁴ traders on the long-term market and 11 on the short-term market. All short-term market traders participated in the long-term market as well. On average, the limit orders of 260 shares were made and traded per period on our long-term market and 410 on the short-term market. During each period, 40% of traders were active on average on the long-term market and two-thirds on the short-term market. The numbers of active traders and trading activity are summarized on the graphs that follow.

⁴Originally, over 50 participants registered in our experiment. However, some of them placed limit orders only during the first period and then stopped; others did not place any orders at all. The remaining 40 participants kept trading throughout the entire experiment.

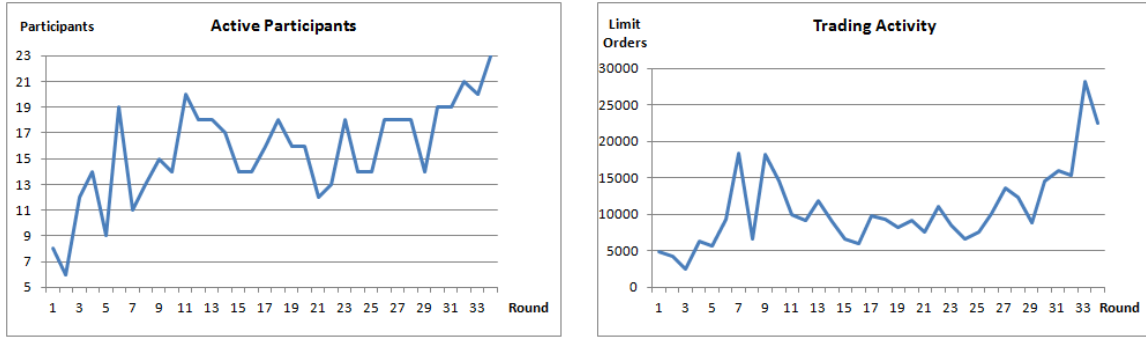


Figure 4. The number of active participants and trading activity on our long-term prediction market.

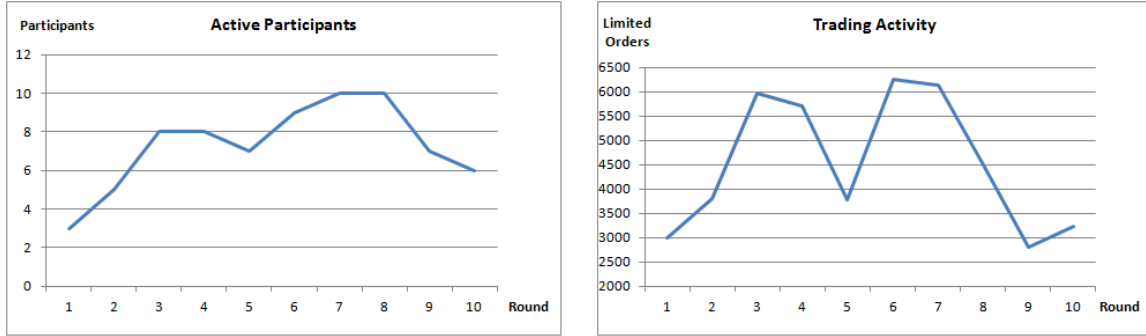


Figure 5. The number of active participants and trading activity on our short-term prediction market.

We do not observe any significant differences in trading activity between the two markets. On the long-term market, we experienced a slight increase in trading activity and in the number of active traders towards the closing period, while activity on the short-term market was slightly decreasing, but the behavior is not significantly different. Call timings — once a day on the long-term market, and twice a day on the short-term market — do not seem to drive any differences in traders' behavior.

Result 1: The short-term market generated more accurate predictions than the long-term market.

Below, we show the distribution of prices generated by both the long-term and short-term prediction markets. Simple prices do not seem to be an optimal measure of the likelihood of individual events. Therefore, similarly to, for example, Chen and Plott (2002), we used volume-weighted average price.

The volume-weighted average price (VWAP) is the ratio of the value traded to total volume traded over a particular time horizon. This way more weight is placed on trades with higher volumes, and we decrease the impact of trades with a few securities

for unreasonable prices, which might occur accidentally. If a price of some security is rather constant over some time and then suddenly during one period the price changes significantly because of a few shares traded, this "accidental" change would be mitigated by using VWAP. We used different time horizons for this purpose. We computed VWAP for the last trading day, for the last trading week, and for the last 50% of trades (as did Chen and Plott 2002).

The VWAP showed to be robust with respect to the choice of a time horizon. The differences in the resulting distributions were rather minor. Hence, we provide the graphs for only one option. In the case of the long-term market, the VWAP is computed over the last week of the market (seven trading days, periods 28-34). In the case of the short-term market, we show the graph for VWAP computed over the last 50% of realized trades, the same as Chen and Plott (2002) for their markets, which were open for one week.

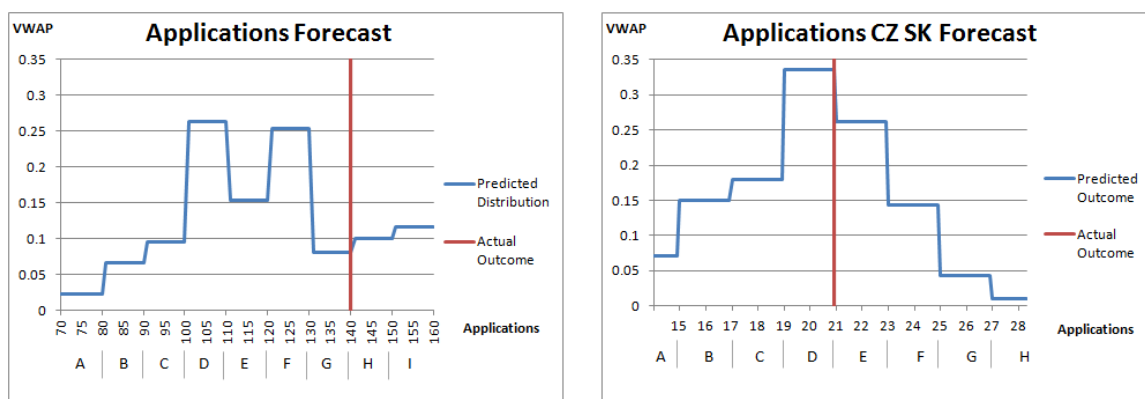


Figure 6. The long-term prediction market with VWAP over the last seven trading days on the left and the short-term market with VWAP over the last 50% of trades on the right. The red vertical line depicts the actual number of applications.

Visual inspection shows that the short-term market performed really well. Based on the distribution generated on this market, event D corresponding to 20-21 applications was identified as the most likely outcome. The actual number of applications received was indeed 21.

In the long-term prediction market, the forecast was less accurate. The distribution generated by the traders' activity on the market suggests, that events D and F are the most likely to occur. These events correspond to 101-110 and 121-130 received applications. However, the actual number was as high as 140 applications. From Figure 6

above, we can see that the distribution of probabilities derived from market prices is multimodal. In particular, we have two favorites on our market. This can simply mean that there were two groups of traders with different expectations. The fact that the prices on the market can imply multinodal probabilistic distribution is illustrated, for example, in Berg, Geweke, and Rietz (2010) for the Iowa prediction market data.

The accuracy of predictions is no doubt of great interest. However, the more important issue here is whether the markets performed well in the sense that they have the capacity to reflect the aggregated knowledge of all the traders, or whether they provide an improvement over more traditional forecasting methods. The choice of proper benchmark and assessment of markets' performance are discussed in the next section.

Result 2: Our prediction markets worked well.

In this section, we argue that our markets worked well, and the distribution of prices sufficiently reflected the collective knowledge of the traders. The purpose of prediction markets is to aggregate existing information. It might happen that traders do not possess any significant knowledge, and if there is not much to aggregate, the markets cannot generate accurate predictions.

At this point, we have to choose a proper benchmark against which we can determine how well the information was aggregated. In laboratory experiments, the information distributed among traders is known, and hence, it is possible to measure the level of information aggregation on the market. Unlike in laboratory experiments, in our case the total information spread among traders is not known, and hence, it is difficult to assess whether all the information in the hands of individual traders was aggregated or not.

We decided to use historical data for the purpose of creating the benchmark. We look at the number of applications during the four years before our market was organized. In particular, we look at the number of total applications and the number of applications from Czech and Slovak students for years 2005/2006; 2006/2007; 2007/2008; and 2008/2009.

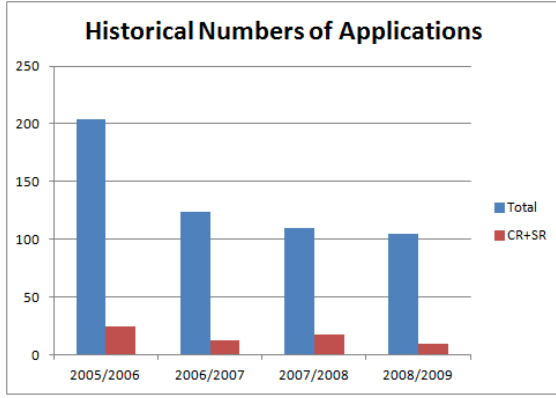


Figure 7. The number of applications received during the four years preceding our markets.

If we look at the data from previous years and consider the decreasing trend, our "best guess" of the number of applications would be 100 in total and 10^5 for the applications from Czech and Slovak students. Comparing this best guess to the actual number of applications (140 in total and 21 for Czech and Slovak applications) and to predictions of our markets (101-110 or 121-130 in total and 20-21 for Czech and Slovak applications) suggests that our markets provided a more accurate prediction than this best guess.

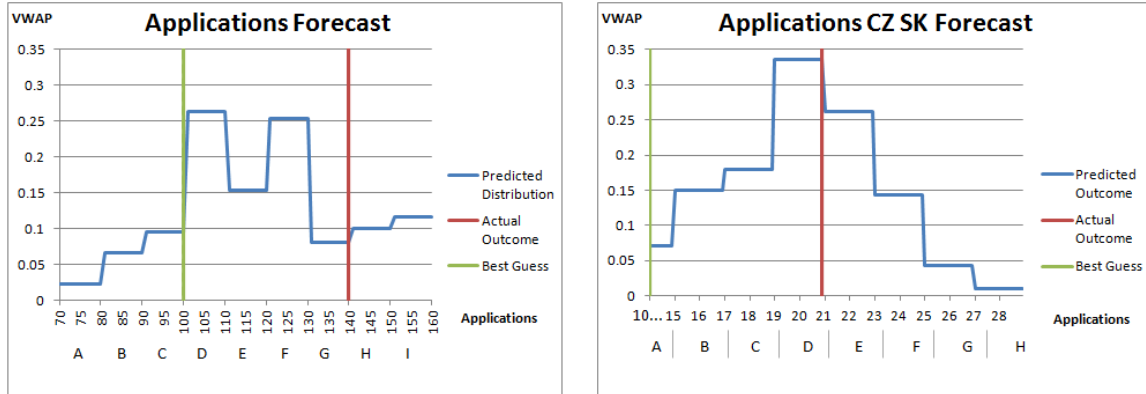


Figure 8. Long-term (on the left) and short-term (on the right) prediction markets with VWAP and "Best guess".

Result 3: The sum of prices on the long-term market is greater than one.

In this section, we look at the sum of prices of all events. On the prediction market, market prices can be interpreted as the probabilities of individual events. Therefore, the prices of all possible events or outcomes should naturally add up to one. In general, the efficiency of the prediction market requires that the sum of all prices equals one. If this

⁵Alternatively, we could choose the average value of received applications, which would actually work very well in the case of our long-term market. However, if we look at how much the number of applications differ every year, we definitely should take the observed trend into account and not rely on simple averages.

is not the case, then arbitrage possibilities might exist and traders could gain risk-less profit.

On our markets, traders were allowed to buy/sell any number of market portfolios for the price of 1 Virtual Dollar. Should the sum of prices on the market be consistently greater than one, traders could make a profit by buying market portfolios from a market maker and selling the assets on the market. Similarly, if the sum of prices was consistently lower than one, traders could make a profit by buying assets on the market and selling market portfolios to a market maker.

We first establish the fact that on our long-term prediction market the sum of prices is greater than one. Then, we provide possible explanations for this fact, and at the end of this section, we discuss the situation on our short-term prediction market. In the next section (Observation 1), we explain why theoretical arbitrage possibilities could not be exploited on our experimental markets. Our null-hypothesis is as follows:

H_0 : The sum of prices on the long-term market is equal to 1.

Based on the results of a t-test, we can reject this hypothesis at the 5% level of significance (with p-value equal to 0.011) in favor of the alternative hypothesis that the sum of prices is higher than one. To test this hypothesis, we use data from the first 34 out of 35 periods. The reason for excluding the last period is the fact that only a very limited number of trade orders were made, and market prices could not be determined for six out of nine events.

We also used a non-parametric Wilcoxon rank sum test and got qualitatively similar results. The null-hypothesis can only be rejected at the 10% level of significance, which is most likely caused by a relatively small number of observations that we have. Moreover, in some periods (9 out of 34), some of the assets were not traded, and therefore, their price could not be determined. As a result, the sum of the prices did not include their value and hence is understated.

The fact that the sum of prices is greater than one can be observed also from the graph that follows. While the sum of prices was oscillating, during most periods, it was

greater than one with a mean value equal to 1.15.

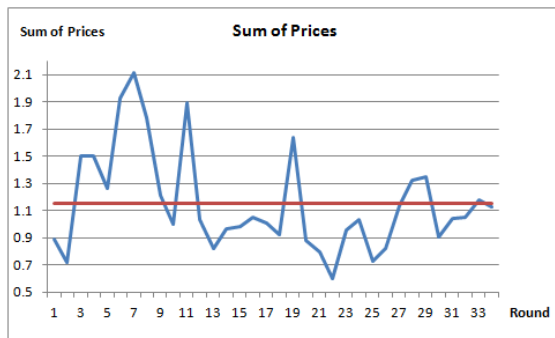


Figure 9. The sum of prices on our long-term prediction market during 34 periods: Mean value, which is equal to 1.15, is depicted by the red horizontal line.

Now that we established the fact that the sum of prices on our long-term prediction market was greater than one, we discuss a possible explanation of this phenomenon. We think that this market inefficiency was (at least partly) caused by the incentivization scheme on our long-term prediction market. This market was organized in such a way that only the three most successful traders earned prizes. In order to increase the chances of profit, traders engaged in a strategy where they were trying to hold as many assets of their favorite event as possible. This motivation pushed prices a little above their true value, and as a result, the sum of all prices exceeded the value of one. This conjecture could be experimentally tested by conducting two prediction markets, which would differ only in the payment scheme. We leave this for further research.

To some extent, we can test this conjecture with the use of our short-term prediction market. In the short-term market, traders were not endowed with any experimental money, and they were not competing for a few big prizes. When using their own money, traders were more motivated to trade assets for their true value. In the case of this short-term market, the null-hypothesis that the sum of prices equals one cannot be rejected. However, since our long-term and short-term prediction markets differ in many criteria, we cannot draw any definite conclusions here.

Result 4: Activity of traders does not affect their profit.

In this section, we test the hypothesis that more active traders earn higher profit. To measure the level of traders' activity, we use a simple measure of the average number of shares traded per period. We use traded shares instead of the number of limit orders placed to avoid the potential problem with most active traders being those who place many unreasonable orders.

In the short-term market, the difference in activity across traders is rather negligible. Therefore, we concentrate our attention on the long-term prediction market, where the difference in the activity of traders is more apparent. The average number of traded shares per period across individual participants is depicted in the figure below. The median value is 111. We divided all participants into two groups: more active traders (with an average of traded shares above median) and less active traders (with an average of traded shares below median value).

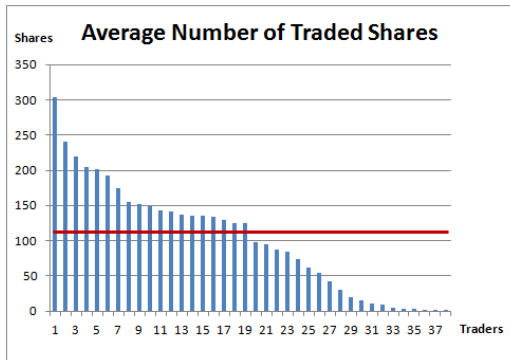


Figure 10. Average number of traded shares per trader and period. The median, which is equal to 111, is depicted by the red horizontal line.

Our initial expectation was that more active traders should be more successful and earn on average higher profit than those who only rarely take some action. However, our analysis leads to the opposite result. The group of more active traders earned on average 30% lower profit than the group of traders who were less active. This result is not statistically significant; nevertheless, we provide our explanation for this interesting observation. Our null-hypothesis is as follows:

H_0 : *The average earning of more active players is the same as the average earning of less active traders.*

Based on the results of a Wilcoxon rank-sum test, we cannot reject the null-hypothesis at 10% level of significance. We conjecture that the reason for this result is as follows. In our market, several traders chose an aggressive strategy. At the very beginning of the market, these participants chose one or two events, and during the entire market, they tried to buy as many shares of these assets as possible.

However, none of these chosen assets turned out to be a winning event. Therefore, this group of the most active traders ended up with extremely low profits. On the other hand, those participants, who were less active and held evenly distributed shares of each event ended up with moderate earnings.

Only two faculty members actively participated on our market. Therefore we were not able to analyze if students and faculty members behave differently, or if faculty members earned on average higher profits on the account of having supposedly more accurate information about interest in the CERGE-EI Ph.D. program.

Observation 1: The sum of the prices was greater than one, but arbitrage possibilities did not exist.

In some periods, prices summed to a number greater than one, which was the value of the portfolio that consists of one contract for each event. We observed this phenomenon mostly (though not exclusively) in the early periods of both markets probably due to the fact that traders were just learning about the market value of contracts, and to make sure that their orders would be matched, they placed very high buy orders.

In theory, this violated the no-arbitrage conditions. However, to take advantage of the arbitrage conditions, individuals would need to know in advance that the prices will add up to a number greater than one, and they would have to sell one piece of each asset in order to exploit this opportunity. Since the prices were fluctuating, this was very unlikely to happen. Traders were placing limit orders without knowing what the market price will be, and only after the period was closed, were orders matched and prices determined. In theory, it could happen that a trader chose such limit orders that he would sell shares of each event and make a profit. However, the oscillating sum of prices around the value of

one clearly shows that this could not have been done systematically.

Moreover, should some traders try to profit on buying portfolios from market makers and selling them on the market, this would push prices down, and theoretical arbitrage opportunities would disappear. Thus, although violations of theoretical arbitrage conditions were observed in the experiments, there were actually no practical arbitrage opportunities.

Observation 2: Non-winning events were overpriced.

We look at the last week of trading on our long-term market and compare prices with the actual number of received applications in a given period. In theory, prices of all events corresponding to a lower number of applications than was received on a given day (trading period) should be essentially zero. However, this is not what we observe in our market.

To illustrate the situation on our market, we provide two graphs for periods 32 and 34 of the long-term market. The red vertical line corresponds to the actual number of applications received on a particular day. If the prices on the market reflect all the information available to the traders, then the prices of all events to the left of the red vertical line should be zero.

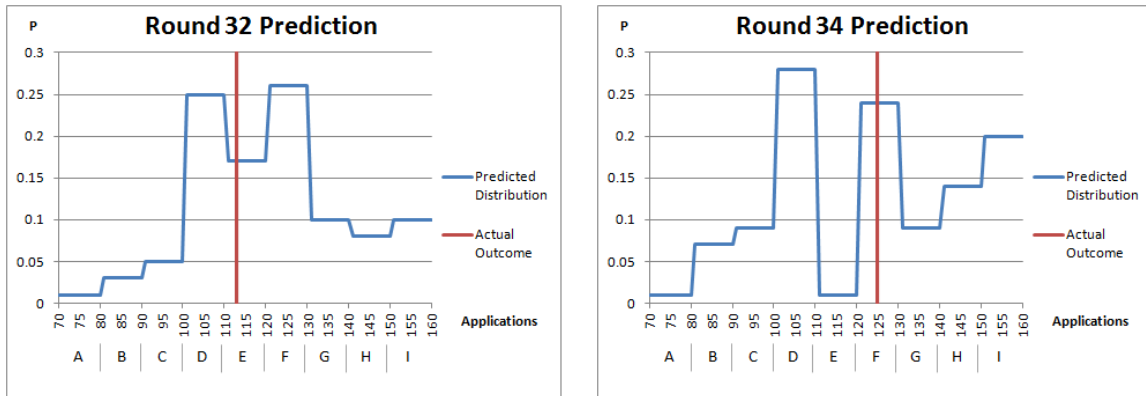


Figure 11. The distribution of prices on the long-term prediction market in period 32 on the left and in period 34 on the right: The red vertical line depicts the actual number of received applications in a given period.

From the graphs it is clear that shares corresponding to events A, B, C, and especially D are significantly overpriced. A detailed inspection of the data of the individual buy

orders shows that the most likely explanation is that traders did not try to get all the information about the current number of received applications. This information was available in the student office (as well as the information about the applications received for past years), but the numbers were not made public, traders had to ask for them.

Here the question of the optimal role of a market maker on small internal markets arises. It seems that a market maker might want to encourage participants to gather all possible pieces of information or maybe directly provide the necessary facts, especially, if there is a reason to assume that not enough insiders would participate in the market.

We also analyzed our short-term prediction market designed to forecast the number of applications from the CR and SR only. In this short-term prediction market we did not experience the problem with overpriced securities corresponding to events that will not occur for sure. As the number of received applications increased over time, the price of shares tied to events A and B dropped, and eventually, these shares were not traded at all because there were no buy orders placed.

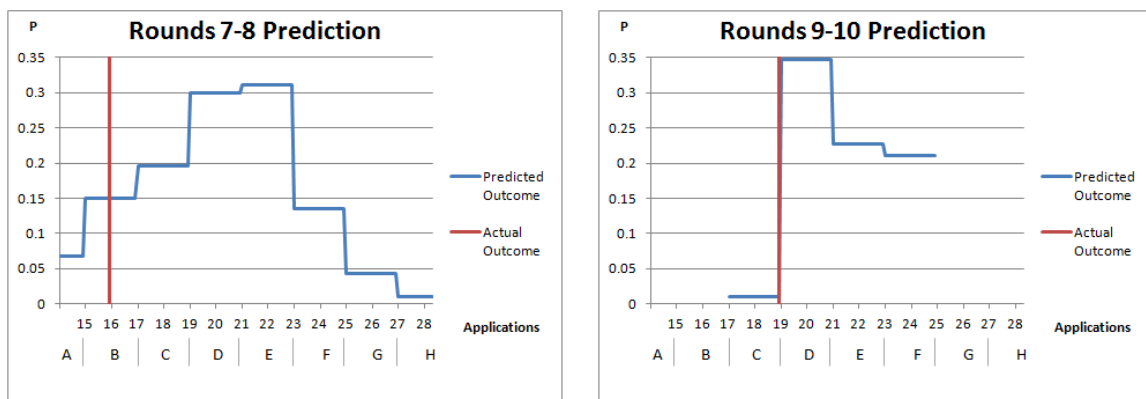


Figure 12. The distribution of prices on the short-term prediction market in periods 7-8 on the left and in periods 9-10 on the right: The red vertical line depicts the actual number of received applications in a given period.

Observation 3: Learning was faster in our short-term market.

In this section, we look closely at the last few trading days on both markets. As we discussed in Observation 2, non-winning events were overpriced. In other words, during certain periods towards the end of the market, subjects were buying events that were tied to a rather low number of applications, while the actual number of received applications was already higher during that time.

Long-term prediction market

Events corresponding to the low number of applications were overpriced in our long-term market. We think that this happened because of two reasons. First, the information about the actual number of applications got to traders slowly. This information was not publicly available, traders had to contact our student office to get this information, so obtaining the information was costly.

The second reason appears to be the incentivization scheme on our long-term market. As we already have mentioned, only the three most successful traders were awarded with a prize, other participants got nothing. Therefore, to have a chance of winning some prize, traders used their a priori beliefs to identify the most likely outcome, and they were placing buy orders on the corresponding security from the beginning of the market.

As the number of received application kept increasing, it was a sign that events A, B, C, and D are unlikely to occur. But traders who concentrated on buying corresponding securities already invested almost all of their money into these securities, and it was too late to change their strategy. Hence, these traders ignored the signals about the increasing number of applications, and the prices of these events were higher than their objective probabilities. Only after it was clear that the number of received applications exceeded the numbers corresponding to events A, B, C and D, did our subjects slowly stop placing orders on these events.

To support this argument, we look at those traders, who concentrated on buying shares of securities tied to a low number of applications. In particular, we analyzed the behavior of traders buying shares of events A-D. There were 9 such traders. They can easily be identified because, on average, these traders held over 2000 shares of one or two events A-D and on average less than 200 shares of all other events.

These traders were very active during the experiment until the last few periods. They were placing high-volume buy orders on events A-D until period 32, when it became clear that these events will not be winning. In period 33, suddenly half of these traders stopped placing orders completely. However, they were still active on the short-term prediction market. This clearly shows that these traders hoped for a low number of applications

until the very last moment and kept the prices of corresponding events high despite all signals suggesting that the number of applications would be much higher.

Those 5 traders who kept placing buy orders on events A-D even after period 33 were most likely not aware of the current number of received applications. Due to these few traders, market prices were kept relatively high until the very end of the market.

Short-term prediction market

Our short-term prediction market was much more flexible, and the signals about the increasing number of received applications were reflected in prices much faster. In this market, traders were not competing against each other in a tournament. Payoffs were determined based on the value of the portfolio held after the market was closed. Therefore, all individuals were trying to increase the number of the securities tied to the most likely event. This simply means that immediately after receiving the signal about the increasing number of received applications, traders updated their beliefs and concentrated on obtaining assets corresponding to the higher numbers of applications.

This observation suggests that the system of incentives has a significant impact on how fast the new pieces of information get to the market through traders' actions. For the reasons mentioned above, there are no significant trends in the sequences of predictions in the case of our long-term market. In other words, the distribution of prices does not change significantly during the last days even though new information was available every day. In the case of our short-term market, we observe a trend in prices that is consistent with the expectations of the future increase of the number of applications.

Observation 4: The number of buy offers exceeded the number of sell offers.

Our result is opposite to that of Chen and Plott (2002), where the authors report that the number of sell offers exceeded the number of buy offers. On our markets, the majority of placed orders were reasonable. Orders to buy at the price equal to zero or close to zero and orders to sell for a price equal or close to one account for only approximately 5% of all placed orders.

In our long-term market, the number of buy offers is higher than the number of sell

offers. We suggest that this result is caused by the tournament incentive scheme in our market (as opposed to the standard reward structure employed in Chen and Plott 2002). As we discussed before in the section on the design of our market, if only the most successful traders get a prize, they are highly motivated to buy as many shares of the most likely events as possible. This claim accords with evidence presented in James and Isaac (2000), where the authors show that in tournament, traders often employ a risky, all-or-nothing strategy.

Traders in our market were trying to sell all shares of events which they considered not likely to win and tried to buy the highest possible amount of their favorites. There were 9 events in our long-term prediction market, and traders were endowed with cash and 200 portfolios. Even if a trader sells all of her shares of non-winning events, she only sells up to 1600 shares (8 events not likely to win with 200 shares of each). The number of shares of a favorite event⁶ was significantly higher than that for the majority of active traders (around 2500-3000 shares). This fact sheds light on why we observe more buy than sell orders on our market. We observe similar behavior on our short-term market. However, here the difference between the number of buy and sell offers is very small.

Observation 5: The success of the internal prediction market appears to be dependent on the nature of the institution and the market itself.

Compared to internal prediction markets running within corporate businesses such as Hewlett-Packard, Siemens and others, our long-term prediction market was less successful in predicting the outcome of our interest. In this particular case, a poll, where the individuals with the most accurate guesses get a prize, might be a faster, simpler and probably equally precise way to gather the dispersed pieces of information. On the other hand, the short-term prediction market with different implementation features worked very well and identified the winning event successfully.

⁶Most traders were investing into buying shares for one or two events.

Observation 6: The incentive scheme in our long-term internal prediction market might have hindered the process of information aggregation.

The ability of the prediction market to form an accurate prediction is based on the process of learning about other people's beliefs by observing their actions. In other words, individuals enter the market with their subjective beliefs and over the time, they update these beliefs based on observing the behavior on the market.

In our long-term market, only three of the most successful traders got a prize. Hence, our incentive scheme supported the type of behavior, where the traders bet all the money on their a priori black horse and stick to this event in order to buy as many underlying assets as possible. Even if they later observed that other traders concentrated on different events, it is usually too late to switch to a different event, and hence, the learning process is not reflected in the prices.

3.5 Optimal Market Design

In this section, we discuss the necessary conditions for an internal small-scale prediction market to work well. Our findings are based on the existing prediction market literature and our experience from the CERGE-EI internal prediction market.

- Payment scheme: Participants with relevant information have to be motivated to trade actively based on their information. This requires an appropriate payment scheme. Our experimental market suggests that the information is more likely to be aggregated if all traders are paid a "small" profit according to their performance (buying for low and selling for a high price or buying the winning events cheaply) compared to the system in which only a few of the most successful traders get the "big" prize.

The performance of an internal prediction market also seems to be better if participants are required to (at least partly) invest their own money. Our results accord with those cited in James and Isaac (2000). James and Isaac (2000) argue that a

tournament can cause distorted market performance and divergence from intrinsic value pricing.

- Market design: A proper payoff system discussed in the previous paragraph also helps to minimize the problem with insufficient liquidity that is common for small scale markets.
- Real vs experimental money: Using real money seems to have two effects. First, participants are more careful when making trading decisions (less risky behavior, we observe more diversification in trading), and traders have stronger incentives to participate actively in the market. Another advantage is that more informed traders are attracted to the market if they can lose their own money, and hence, there would be more information present in the system.
- The frequency of market clearing: The frequency of market clearing has to be chosen with respect to the duration of the market. The shorter the market the more frequent the market clearing should be. The reason is that sometimes the start of the market is slow. It might be the case that there is only a little information available at the beginning of the market, and traders need a few rounds to learn about other traders' beliefs.
- The open question remains — what is the optimal number of participants or the optimal information size relative to the market that might be required for effective information aggregation to take place? In particular, we do not want to miss a person with much relevant information; but on the other hand, we do not want to include too many participants without any substantial information.

3.6 Prediction Market Counterfactuals

In this section, we try to analyze what would happen if our prediction market was organized as a betting system. In light of the key characteristics of both prediction and

betting markets, a prediction market seems to be the proper choice of the market mechanism in this setting. We were able to provide sufficient support to all traders, and hence, the complexity of the prediction market was not an issue.

Should our internal prediction market for applications be organized as a betting system, we would definitely observe a significant proportion of late trading. Participants would have strong incentives to try and extract as much information as possible before they place their bets. As a result, no information would be aggregated in the early stages of the market, and the predictions would likely be much less accurate. Even if early bets were placed, the betting system is much more rigid compared to a prediction market, and therefore traders would not be able to translate their updated private beliefs into market prices (market odds) so promptly.

On the other hand, a betting market environment would most likely eliminate the problem of trading shares of not-winning events. In theory, the prices of all shares corresponding to a lower number of applications than was received on a given day (trading period) should be essentially zero. However, in our market, we observed that a significant amount of these shares were traded. As a result, the prices of these shares were much higher than their real value.

We already discussed that most likely this phenomenon is caused by traders not getting the new information about the number of received applications. Another reason could be the existence of a market portfolio. A market portfolio consists of one share for each of the nine events. If traders needed cash, they had the option of selling one or more market portfolios to the market administrator at any point. To be able to do so, traders needed to hold one share of each event. If traders lacked some shares of not-winning events, they needed to buy them (for low but non-zero prices) for this purpose. As a result, the market prices of shares tied to not-winning events were too high. Then, the whole (normalized) distribution of prices (interpreted as probabilities) of the events is rather flat compared to the distribution that we would get on a betting market (with the same traders and their private information).

Overall, we think that the prediction market was an appropriate choice as the predic-

tion mechanism in our setting. More general, prediction markets seem more appropriate in small-scale internal settings.

3.7 Conclusion

The goal of this research was to take a well-known and experimentally tested information aggregation mechanism and implement it in a real institution for internal use to examine if it could work in this particular setting. The mechanism of our choice was the set of markets for several securities each of which was tied to a subset of possible outcomes.

In particular, we conducted internal experimental prediction markets at CERGE-EI, Prague, to forecast the number of applications received for the 2009/2010 Ph.D. program. The hope was to establish the prediction market as a leading indicator of the recruitment process that potentially could affect activities in the weeks leading up to the deadline and the ensuing selection process.

Numerous potential drawbacks appeared during our experimental markets: for example, a low number of participants (especially in our short-term prediction market) or a low level of activity (many participants traded only occasionally). We tried two different payment schemes and concluded that using one's own money discourages some potential traders. On the other hand, those who did participate were more careful about trading and followed the market more closely, and hence, there was a higher potential for accurate predictions. These are all issues which are not a problem in the case of laboratory experiments but become crucial within institution settings.

Overall, the results of this project show promise. We show that for the performance of the market, design and implementation details matter greatly, especially in the case of internal small scale prediction markets. We discuss optimal market design and conclude that a *properly* designed market has the strong potential to serve as a prediction and internal decision making tool.

3.8 Appendix

In this appendix, we provide the full set of instructions received by our traders. All traders had continuous access to this set of instructions during the entire time of the experiment. We include the instructions in this chapter in order to provide all important details of our market concerning, for example, the type of information available to traders, the type of actions that traders could take, the resolution of ties etc.

INSTRUCTIONS - INTRODUCTION

When you log in to the market you will receive a cash balance of \$200 and 200 market portfolios (i.e. 200 shares of each event). You will have access to all the information about your current share holdings, amount of cash, and the number of portfolios that you own, and you will be able to place orders to buy/sell individual shares/market portfolios. You will also have access to detailed information about previous rounds (orders that you placed, trades that were realized, past market prices, etc.) The information will be available in the following form.

Your remaining Wealth, Disposable Wealth (Wealth minus cost of all active orders), share holdings, and portfolios (number of full sets of all shares):
(A: <= 80 applications; B:81-90; C:91-100; D:101-110; E: 111-120; F: 121-130; G: 131-140; H: 141-150; I: >=151)

Current Round	1
Wealth	160
Disposable Wealth	155
Shares of Event "A"	240
Shares of Event "B"	210
Shares of Event "C"	220
Shares of Event "D"	140
Shares of Event "E"	240
Shares of Event "F"	240
Shares of Event "G"	225
Shares of Event "H"	240
Shares of Event "I"	240
Portfolios	140

Current Round is the number of the round you are playing at the moment, and there will be 37 rounds or trading days (January 21 – February 26, 2010).

Wealth is your cash (200 at the beginning). You can increase your cash by selling market portfolios or shares. If you place some buy orders and they are realized, your cash will decrease by the cost of all realized orders.

Disposable Wealth is Wealth minus the value of all the buy orders that you place. You can keep placing Buy orders as long as your Disposable Wealth stays positive.

Shares of Event "X" is the number of shares of Event X that you currently hold. If you start with 100 shares of Event "X", and you place an order to Sell 10 shares of Event "X", this number will decrease to 90. But if this order is not realized during market clearing, the number of shares of Event "X" will be 100 in the next round.

Portfolios is the number of market portfolios you currently hold.

HOW TO BUY/SELL MARKET PORTFOLIOS

Purchasing a Market Portfolio: When you have a cash balance of at least \$1, you have the option of purchasing one or more market portfolios from the market administrator. Each market portfolio costs \$1, and it consists of one share associated with each outcome: 1 share of Event A, 1 share of Event B, ... , 1 share of Event I. The value of this portfolio is exactly \$1, since there will be exactly one observed event ("winner"), so one share in the market portfolio will pay \$1 and the remaining shares will pay \$0. If you purchase 1 market portfolio, your cash balance will decrease by \$1.

Selling a Market Portfolio: When you have share holdings of at least 1 share of each event, you have the option of selling one or more market portfolios to the market administrator. Each market portfolio costs \$1 so each sold portfolio will bring you \$1 in cash.



Buy/Sell market portfolio:
(Portfolios are bought/sold immediately)

Type	Quantity	
Sell	10	Submit

You can buy/sell Market Portfolios by choosing "Type" (Buy or Sell) and "Quantity" and confirming your order by clicking the button "Submit". In the figure above, you can see the order to Sell 10 Market Portfolios. After clicking the "Submit" button, the number of your Market Portfolios will immediately decrease by 10, and Wealth as well as Disposable Wealth will increase by 10.

HOW TO BUY/SELL SHARES

If you have enough cash (Disposable Wealth) you can place an order to buy shares. Similarly if you hold enough shares you can place an order to sell them. When placing an order you choose: "Event" (A-I), "Type" (Buy or Sell), "Quantity" (how many shares you want to buy or sell), and "**Limit Price**". If you place an order to buy, the Limit Price determines the **maximum price** that you are willing to pay for that share(s). If you place an order to sell, the Limit Price determines the **minimum price** that you are willing to accept for that share(s). In the picture below, you can see an order to buy 30 shares of Event "A" for at most \$0.35.

Submit a new order for this round
(Market will be cleared at midnight)

Event	Type	Quantity	Limit Price	
A	Buy	30	0.35	Submit

If you click the "Submit" button your offer will be recorded and you will see all your orders in the current round in the table called "List of current trading orders".

List of current trading orders					
ID	Event	Type	Quantity	Limit Price	Delete it?
1	A	Buy	30	0.35	Delete it!
2	C	Buy	20	0.25	Delete it!
3	G	Sell	15	0.1	Delete it!

If you want to cancel one of the orders, just click on the "Delete it!" button, and your order will be deleted.

Note: The Quantity of an order is limited to 50. If, for example, you want to buy 65 shares of event A, you can do it by placing two orders (e.g. one for 50 and the other for 15 shares).

HOW ARE PRICES SET (MARKET CLEARING)

Rounds (Trading Days): Each day, at 3 a.m. all orders to buy and sell shares for a particular event will be matched up to determine the trades and associated trading prices for shares. (All orders from your List of current trading orders will be taken into account.) This process is called **market clearing**, and the resulting price at which shares

are bought or sold is called the **market clearing price**.

Arranging Trades: Trades are possible if some of the sell order prices are below some of the buy order prices. The market maker is a computer program that will organize the buy and sell orders and use these to determine a market clearing price. Trades will NOT be arranged for ask prices that are above this level and bid prices that are below this level.

Market Clearing: All transactions will be done at the **same market clearing price**. This will generally be a price at which the number of shares that traders wish to buy is equal to the number of shares that traders wish to sell. In other words, the number of shares with limit sell prices at or below this clearing price is equal to the number of shares with limit buy prices at or above this clearing price. Thus, those who are willing to pay the most will buy from those who are willing to sell for the least, but all trades will be at the same price.

Clearing Prices: The bids and asks for each of the 9 types of shares will be used to determine the market clearing price for that type of share. Thus, there will be 9 clearing prices determined each time the market is cleared. All shares of a particular type that are bought and sold when the market clears will be bought and sold at the same price, i.e. at the clearing price for that type of share.

Resolution of Ties: In some cases, it may not be possible to exactly equalize the numbers of units demanded and offered. For example, if there is one limit order to purchase a single share at 2 cents, and if there are 2 limit orders to sell, each for a single share at 2 cents, then the market clearing price would be 2, but one of the two sell orders cannot be executed. The decision of which order to execute in the event of a tie will be based on a random process, i.e. the computer equivalent of the flip of a fair coin.

Examples of Clearing Price Determination:

Example 1

Trader 1: sell 10 shares of Event A for at least \$0.5 Trader 2: buy 5 shares of Event A

for at most \$0.6 Trader 3: buy 5 shares of Event A for at most \$0.7 The resulting price will be \$0.55, and 10 shares will be sold.

Example 2

Trader 1: sell 10 shares of Event A for at least \$0.5 Trader 2: buy 5 shares of Event A for at most \$0.6 Trader 3: buy 3 shares of Event A for at most \$0.7 The resulting price will be \$0.5, and 8 shares will be sold.

Example 3

Trader 1: sell 8 shares of Event A for at least \$0.5 Trader 2: buy 5 shares of Event A for at most \$0.6 Trader 3: buy 5 shares of Event A for at most \$0.7 The resulting price will be \$0.6, and 8 shares will be sold.

Note that in all three examples the market clearing price can be anything between 0.5 and 0.6. For all prices from this interval the same amount of shares will be sold.

The rules for determining the exact price are as follows:

If the volume of demand (buy) and supply (sell) is the same, then the market clearing price is in the middle of this interval (0.55 in *Example 1*).

If the volume of demand (buy) is smaller than the volume of supply (sell), then the market maker chooses the price from the interval, that is the most favorable for buyers, and the least favorable for the sellers (\$0.5 in *Example 2*).

If the volume of demand (buy) is larger than the volume of supply (sell), then the market maker chooses the price from the interval that is the most favorable for the seller and the least favorable for the buyers (\$0.6 in *Example 3*).

History: You will have access to a full history of your actions all the time. At the bottom of the trading page, you can find the button "History of Prices and Trades". If you click on this button, you will be redirected to a page with the history of all the clearing prices for all past rounds. Also, you can review all the orders that you placed in past rounds, and you will see if they were realized or not.

Bibliography

- Asch, P., B. G. Malkiel, and R. E. Quandt. (1984). Market Efficiency in Racetrack Betting. *Journal of Business* 57(2). pp. 165-751.
- Berg, J. E., R. Forsythe, F. Nelson, and T. A. Rietz. (2008). Results from a Dozen Years of Election Futures Markets Research. *Handbook of Experimental Economics Results*, vol. 1. pp. 742-751.
- Berg, J. E., J. Geweke, and T. A. Rietz. (2010). Memoirs of an Indifferent Trader: Estimating Forecast Distribution from Prediction Markets. *Quantitative Economics* 1(1). pp. 163-186.
- Berg, J. E., F. Nelson, and T. A. Rietz. (2003). Accuracy and Forecast Standard Error of Prediction Markets. *Departments of Accounting, Economics and Finance, University of Iowa, Iowa City, IA, working draft*.
- Berg, J. E., and T. A. Rietz. (2010). Longshots, Overconfidence and Efficiency on the Iowa Electronic Market. *Departments of Accounting, Economics and Finance, University of Iowa, Iowa City, IA, working draft*.
- Berlemann, M., K. Dimitrova, and N. Nenovsky. (2006). Assessing Market Expectations on Exchange Rates and Inflation: A Pilot Forecasting System for Bulgaria. *Current Politics and Economics of Russia, Eastern and Central Europe* 21(3). pp. 263-287.

- Berlemann, M., and F. Nelson. (2005). Forecasting Inflation via Experimental Stock Markets. Some Results from Pilot Markets. *Ifo Working Papers No. 10*.
- Bullen, J. (2009). Parimutuel Information Aggregation Mechanisms with Public Knowledge. *The University of Sydney*.
- Cahlík, T., A. Geršl, M. Hlaváček, and M. Berlemann. (2005). Market Prices as Indicators of Political Events. Evidence from the Experimental Market on the Czech Republic Parliamentary Election in 2002. *Working Papers IES 77, Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies*.
- Camerer, C. F. (1998). Can Asset Markets Be Manipulated? A Field Experiment with Racetrack Betting. *The Journal of Political Economy* 106(3). pp. 457-482.
- Chen, K., and C. R. Plott. (2002). Information Aggregation Mechanisms: Concept, Design and Implementation for a Sales Forecasting Problem. *Working Paper Social Science 1131, California Institute of Technology*.
- Debnath. (2003). Information Incorporation in Online In-game Sports Betting Markets. in *Proceedings of the Fourth Annual ACM Conference on Electronic Commerce*. pp. 258-259.
- Eisenberg, E., and D. Gale. (1959). Consensus of Subjective Probabilities: The Parimutuel Method. *The Annals of Mathematical Statistics* 30(1). pp.165-168.
- Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics* 10(2). pp. 171-178.
- Forsythe, R., F. Neslon, G. R. Neumann, and J. Wright. (1992). Anatomy of an Experimental Political Stock Market. *The American Economic Review* 82(5). pp. 1142-1161.
- Forsythe, R., T. A. Rietz, and T. W. Ross. (1999). Wishes, Expectations and Actions: a Survey on Price Formation in Election Stock Markets. *Journal of Economic Behavior and Organization* 39. pp. 83-110.
- Galton, F. (1907). Vox Populi. *Nature* 75. pp. 450-451.
- Hahn, R. W., and P. C. Tetlock. (2006). Information Markets: A New Way of Mak-

- ing Decisions in the Public and Private Sectors. *AEI-Brookings Joint Center for Regulatory Studies*. Washington, D.C.
- Hanson, R., and R. Oprea. (2009). A Manipulator Can Aid Prediction Market Accuracy. *Economica* 76 (302). pp. 304-314.
- Harrison, G. W., E. Johnson, M. M. McInnes, and E. Rutström. (2005). Risk Aversion and Incentive Effects: Comment. *The American Economic Review* 95(3). pp. 897-901.
- Harrison, G. W., and E. Rutström. (2008). Risk Aversion in the Laboratory. in Cox, J. C. and Harrison, G. W. (eds.), *Risk Aversion in Experiments*, vol. 12 (Bingley, UK: Emerald, Research in Experimental Economics).
- Hayek, F. A. (1945). The Use of Knowledge in Society. *The American Economic Review* 35(4). pp 519-530.
- Holt, C. A., and S. K. Laury. (2002). Risk Aversion and Incentive Effects. *The American Economic Review* 92(5). pp 1644-1655.
- James, D., and R. M. Isaac. (2000). Asset Markets: How They Are Affected by Tournament Incentives for Individuals. *The American Economic Review* 90(4). pp 995-1004.
- Kálovcová, K., and A. Ortmann. (2009). Understanding the Plott-Wit-Yang Paradox. *The Journal of Prediction Markets* 3(3). pp. 33-44.
- Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica* 53 (6). pp. 1315-1336.
- Luckner, S., and C. Weinhardt. (2007). How to Pay Traders in Information Markets: Results from a Field Experiment. *The Journal of Prediction Markets* 1(2). pp. 147-156.
- Manski, C. (2006). Interpreting the Predictions of Prediction Markets. *Economic Letters* 91(3). pp 425-429.
- Nalebuff, B. J., and J. E. Stiglitz. (1983). Prizes and Incentives: Towards a General Theory of Compensation and Competition. *Bell Journal of Economics* 14(1). pp. 21-43.

- Ortner, G. (1997). Forecasting Markets - An Industrial Application: Part I. *Working Paper, TU Vienna*.
- . (1998). Forecasting Markets - An Industrial Application: Part II. *Working Paper, TU Vienna*.
- Ottaviani, M., and P. N. Sorensen. (2006a). Aggregation of Information and Beliefs in Prediction Markets. *Mimeo*.
- . (2006b). Noise, Information, and the Favorite-Longshot Bias. *FRU Working Papers No 2006/04*, University of Copenhagen. Department of Economics. Finance Research Unit.
- . (2006c). The Timing of Parimutuel Bets. *Mimeo*.
- Plott, C. R., and S. Sunder. (1982). Efficiency of Experimental Security Markets with Insider Information: An Application of Rational-Expectations Models. *The Journal of Political Economy* 90(4). pp 663-698.
- . (1988). Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets. *Econometrica* 56(5). pp 1085-1118.
- Plott, C. R., J. Wit, and W. C. Yang. (2003). Parimutuel Betting Markets as Information Aggregation Devices: Experimental Results. *Economic Theory* 22(2). pp 311-351.
- Polgreen, P. M., F. D. Nelson, and G. R. Neumann. (2007). Use of Prediction Markets to Forecast Infectious Disease Activity. *Clinical Infectious Diseases* 44(2). pp.272-279.
- Rhode, P., and K. Strumpf. (2004). Historical Presidential Betting Markets. *The Journal of Economic Perspectives* 18(2). pp 127-141.
- Roust, A. K., and R. C. Plott. (2005). The Design and Testing of Information Aggregation Mechanisms: A Two-Stage Parimutuel IAM. *Social Science Working Paper No. 1245*.
- Sauer, R. D. (1998). The Economics of Wagering Markets. *Journal of Economic Literature* Vol. XXXVI, pp. 2021-2064.
- Smith, V. L. (1962). An Experimental Study of Competitive Market Behavior. *The Journal of Political Economy* 70(2). pp. 111-137.

- Thaler, R. H., and W. T. Ziemba. (1998). Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries. *The Journal of Economic Perspectives* 2(2). pp. 161-174.
- Wolfers, J., and E. Zitzewitz. (2006a). Five Open Questions About Prediction Markets. in *Information Markets: A New Way of Making Decisions in the Public and Private Sectors*, ed. Robert Hahn and Paul Tetlock, AEI-Brookings Joint Center, Washington D.C.
- . (2006b). Prediction Markets in Theory and Practice. *NBER Working Papers* No. 12083.
- . (2004). Prediction Markets. *Journal of Economic Perspectives* 18(2). pp 107-126.
- . (2005). Interpreting Prediction Market Prices as Probabilities. *Mimeo, University of Pennsylvania*.
- Wurtz, A. H. (1997). A universal upper bound on power functions. *UNSW Discussion Paper 97/17, Sydney, Australia, University of New South Wales*.