Continuous Market Engineering

Focusing Agent Behavior, Interfaces, and Auxiliary Services

Zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der Fakultät für

Wirtschaftswissenschaften

am Karlsruher Institut für Technologie (KIT)

genehmigte

DISSERTATION

von

Dipl.-Wi.-Ing. Tobias Theodor Kranz

Tag der mündlichen Prüfung: 10. Februar 2015

Referent: Prof. Dr. Christof Weinhardt

Korreferent: Prof. Dr. Orestis Terzidis

Karlsruhe, 2015

Abstract

Electronic markets, driven by the rise of the Internet and the establishment of the world wide web, spread out amongst business entities as well as private individuals. Albeit numerous approaches on designing and developing electronic markets exist, e.g., Roth's *Market Design* or Weinhardt et al.'s *Market Engineering*, a unified approach targeting market development, redesign, and refinement has been lacking.

The work at hand studies the potential of continuously improving electronic markets from a market provider's perspective. It comprises five experiments conducted on different instances of prediction markets targeted to forecast macroeconomic indicators as well as political outcomes. Thereby, the experiments' design focuses on the three distinct aspects of *Agent Behavior*, *Interfaces*, and *Auxiliary Services* in electronic markets.

First, behavioral aspects of market participants (agents) are linked with both, the quality of their trading decisions and trading behavior, by combining trade data with replies of a specifically compiled questionnaire. Hereby, traders' market predispositions, i.e., their aptitude for trading, can be assessed ex ante and might serve as a measure to counsel traders. Second, in a political stock market, traders' political preferences are deduced from their trading behavior. Interestingly, a simple model considering the agents' behavior is yet sufficient to explain participants' political preferences correctly to a high extent. Third, a comparison is drawn of participants' trading performance and trading behavior between a web interface and a similar mobile application. Results show that albeit mobile interfaces are accepted by participants, trading conducted via the provided web interfaces is more successful in a pecuniary sense and provides higher predictive power concerning the predicted event. Fourth, the trading interface's impact on the *Disposition Effect* (i.e., the disposition to hold losing stocks too long whilst selling winning stocks too early) is explored. To this end, two trading interface modifications are evaluated: a trend indicator arrow, reflecting the development of traders' portfolio value, and textual advice about the disposition effect. Results reveal that increasing transparency of traders' portfolio value increases the disposition effect, whereas textual advice does not seem to be apt in decreasing the effect's strength. Fifth, since prediction markets are an established mechanism to acquire quantitative information, an endeavor is made to also assess qualitative information via auxiliary services from dispersed agents, i.e., the crowd. Therefore, external as well as different kinds of internal survey systems are evaluated. The gained research insights comprise a better understanding of trading behavior in electronic markets, and thus, unveil potential ways for market engineers to continuously improve electronic markets.

Acknowledgements

Foremost I would like to thank Prof. Dr. Christof Weinhardt for the opportunity to pursue my dissertation under his supervision. Moreover, I would like to thank my co-advisor Prof. Dr. Orestis Terzidis and the other members of the thesis committee, Prof. Dr. Thomas Setzer and Prof. Dr. Petra Nieken for their helpful comments, their time spent examining the work at hand, as well as the thorough and insightful discussion thereof.

In a fabulous team it is hard to emphasize individuals. Therefore, I would like to thank all colleagues at IISM, FZI, and KSRI for their support and the great working atmosphere! Yet, some gave exceptional support, beyond work, or just in a special way and thus need to be explicitly mentioned. I am deeply indebted to Dr. Florian Teschner for his support, advice, and feedback. For proofreading my work, I deeply thank Anuja Hariharan, Margeret Hall, Felix Fritz, and Dr. Timm Teubner. I would like to thank Jadon Herrmann, Philipp Roüast, and Tim Straub for making day-to-day operations a pleasure. For discussing hot topics in a relaxing manner I would like to thank Prof. Dr. Lukas Wiewiorra, Dr. Philipp Ströhle, Niklas Horstmann, and Daniel Schnurr. Moreover, I would like to thank Dr. Stephan Meyer for making me believe in coincidence and other highlights. From time to time, when perseverance waned, a bitsy hint might help to see things different. Therefore, I would like to thank Prof. Dr. Jan Krämer and Prof. Dr. Henner Gimpel for sharing their insights with me. Last, I would like to thank Marius Müller and Felix Fritz for frequent finance chats and preservation of FiRe traditions.

Throughout life people come and go, but some are so inspiring that they slightly nudge your life towards a different path. Therefore, I would like to thank Adrian Mitter, Anna-Lisa Seidel, Aria Neshat-Doust, Assen Tarlov, Christine Nelte, the people of Gruppe 4 and Group54, 'Herbert', Jürgen Heinrich, Lubomir Nikolov, Naji Sayhi, Said Hasan Hashemi, Stefan Roeling, Werner Dietz, and Zlatko Naumovski for making their acquaintance. I also thank Jan Steinbach for introducing me to Karlsruhe and accompanying me – and your palm tree. Furthermore, I would like to thank Thorsten Schmidt, Niko Lukesch, Kai Schomaker, Peter Schmidt, Christian Reuter, Stefan Steinbrück, Miloš Kostić, Stevan Đorđević, Philipp Speiser, and Petar Trifonov for your intense and enduring friendship. For bestowing me lifetime by using her power wisely, I thank Ms. Manthey. I owe you a lot!

Finally, I would like to thank my beloved family; my parents Carmen and Klaus for their constant backing and love throughout my life and my sister Valeska for being an inspiring example in many ways.

Karlsruhe, February 2015 Tobias T. Kranz

Contents

Part I

Introduction

Chapter 1

Introduction

"If the world was stable, there would be no need to change business operations and methods, nor to understand what has changed and what works well. However, firms operate in dynamic environments, not stable ones."

ABBIE GRIFFIN, 1997

1.1 Motivation

MARKETS are very efficient and effective mechanisms to exchange goods and services. From classical market places where physical goods are exchanged, through stock markets, up to emissions trading, markets are used. Individuals in modern societies rely on numerous markets in their daily life – often even unknowingly. Business-to-business (B2B) markets, for example, belong to a family of markets most people never interact with directly. Nevertheless, when you switch on the lights in the morning, the electricity was probably traded on an exchange.[1] When you buy goods in retail stores, the price is most likely influenced by wholesale market prices. When you book a flight, the price depends inter alia on foreign exchange markets, as they determine currency prices

[1] In case of Germany energy is often traded on the European Energy eXchange (EEX) in Leipzig (cf. http://www.eex.com/).

and thus affect jet fuel prices. Participants in B2B markets are often professionals with a certain trading experience. With the rise of the Internet and the establishment of the world wide web[2] (WWW) in the 1990's, electronic markets started to attract business entities in addition to private individuals. New forms of electronic markets emerged, often separated into business-to-business (B2B) e-commerce and business-to-customer (B2C) e-commerce.[3] Historically, electronic markets are a rather new way of implementing market mechanisms and have a couple of advantages. They hardly need physical space, transaction costs are often strikingly reduced, and market rules can be changed and monitored centrally. Besides the use-cases mentioned above, markets are also used to trade information or even expectations about future events as will be subsequently shown.

The design of markets plays a pivotal role in relation to 'how well' markets function. Generally, a market's objective influences its requirements, which in turn determines the market's resulting design. This comprises, inter alia, the design of the transaction object, the market microstructure, or, in case of electronic markets, the market's implementation and thus its graphical user interface. Consequences emerging from certain design decisions are a highly discussed issue in economic literature. A seminal article concerning this topic is Roth's (2002) work on *Market Design*, wherein he proposes to transition the planning process of markets from a rather conceptual design approach into a more rigid engineering approach based on theoretical insights. In his 2008 follow-up article he defines three preconditions for properly functioning markets: First, markets have to provide *thickness*, which means *"[...] to attract a large enough proportion of the potential participants [...]"* (Roth, 2008). Second, overcoming *congestion*, which can be evoked by thickness; i. e., providing a sufficient amount of adequate alternative transactions. Third, markets have to be *safe* and easy to use in order to be attractive for participants and thus preventing potential participants from circumventing the market. The idea of market design is revisited and taken further by Weinhardt et al. (2003), resulting in the concept of *Market Engineering* as most recently described by Gimpel et al. (2008). Subsequently, the Market Engineering concept was applied to the development of numerous markets; inter alia a sports prediction market for the FIFA Worldcup 2006 called STOCCER (Luckner et al., 2005) or a

[2]An interesting background report on the ideas and concepts of the WWW can be found Berners-Lee and Fischetti (2000).

[3]Teo and Ranganathan (2004) define business-to-business (B2B) e-commerce as *"the buying and selling of products and services among businesses"* and business-to-customer (B2C) e-commerce as *"the sale of products and services to individuals"*.
A slightly more precise definition of B2B e-commerce can be found in Lucking-Reiley and Spulber (2001) and reads: "The popular phrase 'B2B e-commerce' refers to the substitution of computer data processing and Internet communications for labor services in the production of economic transactions."

multi-attributed combinatorial exchange called MACE (Schnizler et al., 2004; Schnizler, 2008). From another point of view, slightly restructuring Roth's (2008) prerequisites, even the best designed market will only be successful if (i) it is accepted by the participants and (ii) participants behave as expected by the market designer, i. e., assumptions made about market participants' behavior hold true.[4] Therefore, it is indispensable to develop a deeper understanding of market participants, since ultimately, participant behavior and hence the participants' acceptance of a market is the main determinant for a market's outcome and thus its success. For instance, individual's risk aversion or cognitive abilities have been shown to influence decision making (e. g., Subrahmanyam, 1991; Frederick, 2005). Furthermore, emotions are known to impact individuals economic decision making (e. g., Loewenstein et al., 2001). Hence, this thesis attempts to include these aspects of individual behavior in the consideration of market participants' behavior in electronic markets. Following the typology of *Service Analytics* (Fromm et al., 2012), this approach can be classified as advanced analytics on customer data. Bearing in mind that Market Engineering is not solely the engineering perspective of designing and introducing of markets, but also describes their further development, the aspect of *continuously* tracking and improving a market – from a market provider point-of-view – is up to now under-addressed. Additionally, up to now, economic literature lacks an explicitly documented case of continuous tracking and supporting a market in the context of Market Engineering.

Why, should markets be tracked, supported, and improved in a continuous manner? For what reasons could a fit-and-forget approach for markets be insufficient and eventually fail? At least two views on that issue should underline the necessity for *Continuous Market Engineering*. First, it can be observed that numerous market providers repeatedly changed and extended their markets for intrinsic motivated reasons – for instance, reaching out to a new set of clients by satisfying their needs or implementing new features to enhance customer loyalty. It can further be observed that a range of markets adapted or extended their functionality to improve internal processes. Second, as time goes by things change. Although, admittedly, this reads trivial, it leads to the class of external motivated changes. It implies that design decisions drawn once, based on reasonable assumptions at a particular point in time, have a probability to not hold true forever. In other words, assumptions leading to once sensible design decisions can be increasingly violated and even lose their validity due to omitted adaptations to a changing environment, trends in participant behavior, or due to technological causes. Consequently, solely the aim of conserving a

[4]The first point primarily relates to Roth's (2008) second and third prerequisites ('thickness' and 'safe'), whereas the second point concerns all three preconditions.

market state once intended can be a sufficient necessity to continuously track and improve a market – and thus motivating *Continuous Market Engineering*. In contrast to changes in the socio-economic and legal environment, as laws and regulatory standards, changes in individual behavior are not as easy to spot or foresee. Hence, the work at hand stresses the importance of *Continuous* Market Engineering by (i) documenting two continuously engineered markets as well as (ii) developing a deeper understanding of market participant's behavior.

Summing up, stakeholders of a market can benefit from a better understanding of market participants' individual behavior. Since environmental conditions and individual behavior are subject to change, market design decisions have to be monitored and re-evaluated. Markets require continuous adaption and redesign in order to constantly achieve their desired market outcome. The work at hand furthers the understanding of individual trader behavior and demonstrates the power of Continuous Market Engineering by presenting several studies conducted on two play-money prediction markets.

1.2 Research Outline

This thesis stresses the importance of a *continuous* approach to the *Market Engineering* framework (Weinhardt et al., 2003; Gimpel et al., 2008) from a market provider's point of view while focusing on individual trader behavior.[5] Specifically, the Market Engineering framework as presented in Section 2.4 provides guidance for engineering electronic markets. It comprises a specified prescriptive design process model and a toolbox of methods supporting that aim. The research objective of the work at hand is to apply *Continuous Market Engineering* on electronic markets, thus demonstrating its potential and necessity for successful markets. Furthermore, based on the above mentioned framework, a Continuous Market Engineering Process (cf. Section 2.6) is derived from the experiences gained so far. The studies are conducted on two prediction markets presented in Chapter 4 and presented closely following the *Market Engineering Object* (cf. Figure 2.4a).

Personal attributes like risk aversion (RA), cognitive reflection abilities (CRA), and emotion regulation strategies (ERS) are known to influence individual behavior. Up to now, it seems unclear how these personal attributes influence (i) trading behavior (i. e., trader behavior *in* the market) and (ii) decision quality (i. e., quality of trading decisions) in a

[5]The term 'trader behavior' corresponds to 'Agent Behavior' in the market engineering nomenclature (cf. Chapter 2).

prediction market context. Hence, advanced *Service Analytics* (Fromm et al., 2012) are applied on the customer data of a prediction market in order to better understand *Agent Behavior*. This approach delivers comprehensive insights on traders' predisposition for markets. Based on the gained knowledge, it is feasible to substantially improve service experience. This can be achieved by adapting the *Market Structure* to agent's preferences and abilities via personalized tweaks (interface adaption, product choice, etc.). In particular, the following questions are addressed:

> **Research Question 1:** *How do selected personal attributes (RA, CRA, and ERS) influence trading behavior in markets?*

> **Research Question 2:** *How do selected personal attributes (RA, CRA, and ERS) influence decision quality in markets?*

After it has been shown that personal attributes do influence trading behavior, the question arises whether trading behavior can reveal information about the traders themselves. For instance, whether it is possible to 'read a trader's mind' from his actions performed on the market. Therefore, individual trading behavior on a prediction market is analyzed in order to estimate individuals' preferences. Specifically, on a Political Stock Market, the relation between trading behavior and political preferences is examined by answering the following research question:

> **Research Question 3:** *How well can an unobtrusive analysis of trading behavior reveal trader preferences?*

After examining traders' actions on markets from a behavioristic angle, the market providers perspective is taken. It is common sense, that interface design can influence individual decision behavior and thus decision outcome (e. g., Kauffman and Diamond, 1990). With the rise of mobile information systems, the question arises whether and which impact different devices have on decision behavior and decision outcome. Furthermore, the setting in which individuals tend to use mobile devices often differs from stationary settings. Hence, the following research questions is evaluated in this thesis:

> **Research Question 4:** *Are decision behavior and decision outcome affected by the kind of device used?*

Besides interface design, behavioral biases do influence and often harm individual's decision behavior (e. g., Tversky and Kahneman, 1974). An important and well studied behavioral bias especially traders tend to suffer from is the *disposition effect* (cf. Shefrin and Statman, 1985). An endeavor is made to gain insights how differences in trading interfaces influence the disposition effect. First, the effect of providing information about the disposition effect in trading interfaces is examined. Second, the impact of a commonly used interface element of electronic market systems, a trend indicator arrow reflecting a trader's portfolio state, is investigated. Considering the relative strength of behavioral biases in the face of small interface changes, two research questions are addressed:

Research Question 5: *Is providing information about the disposition effect suitable to lower the strength of the disposition effect exhibited by an individual?*

Research Question 6: *Does a trend indicator arrow affect the strength of the disposition effect exhibited by an individual?*

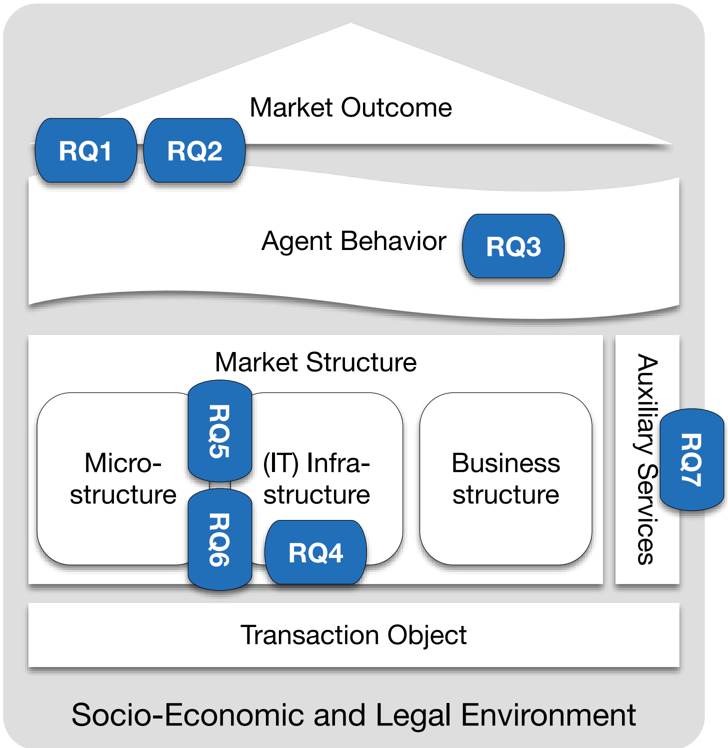
Prediction market operators try to gain insights about future events *from* participants. Market operators may, every once in a while, also have the desire to gather feedback or learn more *about* their participants. Insights gained through such endeavors might help to improve the market and thus raise loyalty as well as attract more customers. Still focusing on the user interface, an attempt is made to find a suitable way to acquire such feedback and additional information *from* and *about* market participants. Although, much research has been conducted on how to design and conduct questionnaires (e. g., Babbie, 1990; Fowler, 2014; Andres, 2012) as well as whom to include in a survey to reach certain goals (e. g., representativeness, response rate), little is known on where exactly – in a technical sense – to conduct surveys. In the specific domain of prediction markets, as well as in the broader domain of online communities, it is largely undecided whether it is beneficial to integrate a survey on the same platform versus conducting the very same survey on an external specialized survey platform. Besides implications on the graphical representation and on the user guidance, this decision impacts development as well as operation costs of the market. Therefore, the final aspect of this thesis poses the following research question:

Research Question 7: *Are integrated surveys more accepted by participants of a prediction market than standalone surveys?*

The aforementioned research questions can be assigned to the components of the *Market Engineering Object* as shown in Figure 1.1.[6] From a Market Engineering perspective,

[6]For an in-depth description of the Market Engineering framework, containing the Market Engineering

research questions 1 and 2 connect *Agent Behavior* with *Market Outcome*, whereas research question 3 solely focuses on *Agent Behavior*. The *(IT) Infrastructure* is subject to research question 4. Research questions 5 and 6 additionally concerns the *Microstructure*. Finally, research question 7 focuses on *Auxiliary Services*.



RQ1: *How do selected personal attributes (RA, CRA, and ERS) influence trading behavior in markets?*
RQ2: *How do selected personal attributes (RA, CRA, and ERS) influence decision quality in markets?*
RQ3: *How well can an unobtrusive analysis of trading behavior reveal trader preferences?*
RQ4: *Are decision behavior and decision outcome affected by the kind of device used?*
RQ5: *Is providing information about the disposition effect suitable to lower the strength of the disposition effect exhibited by an individual?*
RQ6: *Does a trend indicator arrow affect the strength of the disposition effect exhibited by an individual?*
RQ7: *Are integrated surveys more accepted by participants of a prediction market than standalone surveys?*

FIGURE 1.1: *Placement of Research Questions in Market Engineering Object (based on Gimpel et al., 2008)*

Object, see Section 2.4. Note, that some research questions relate to multiple components of the Market Engineering Object. In such cases, a best-fit approach is taken and the research question is assigned to the component, which it relates to most.

1.3 Structure of the Thesis

The work at hand is structured in five parts as shown in Figure 1.2. Part I motivates this thesis, presents the examined research questions, and gives an overview of the structure of this thesis. Part II introduces *Continuous Market Engineering* (see Chapter 2) and presents theoretical foundations on *Prediction Markets* (see Chapter 3) before two instances of prediction market are described (see Chapter 4).

Part III presents selected insights derived from applied Continuous Market Engineering conducted on the two prediction markets described in Chapter 4. Chapter 5 examines how *Agent Behavior* can be analyzed and whether certain characteristics lead to a specific *Market Outcome* (hence addressing research questions 1 and 2).[7] Further analyzing *Agent Behavior*, Chapter 6 describes what a specific trading behavior denotes about a trader's attitude and beliefs (hence addressing research question 3).[8] Chapter 7 analyzes how *Agent Behavior* is influenced by the choice of *(IT) Infrastructure* offers (hence addressing research question 4).[9] Chapter 8 shows how small design changes in the *(IT) Infrastructure* can raise or lower a trader's behavioral bias (hence addressing research questions 5 and 6).[10] Chapter 9 describes how *Auxiliary Services* can be provided that simultaneously improve a trader's platform experience as well as support the market provider's goals (hence addressing research question 7).[11]

Part IV concludes this thesis by stating the contributions made in the context of Continuous Market Engineering, giving a research outlook, and summarizing this work. Appendices are contained in Part V.

[7]Parts of this chapter are based on joint work with Florian Teschner and Christof Weinhardt and have been presented at the *Hawaii International Conference on System Sciences 2014* (Kranz et al., 2014b).

[8]Parts of this chapter are based on a joint article with Florian Teschner, Philipp Roüast, and Christof Weinhardt which has been presented at the *International Conference on e-Society 2014* (Kranz et al., 2014).

[9]Parts of this chapter are based on joint work with Florian Teschner and Christof Weinhardt and have been presented at the *Multikonferenz Wirtschaftsinformatik 2012* (Teschner et al., 2012). Other parts of this chapter are based on a joint article with Florian Teschner and Christof Weinhardt which is published in the *International Journal of E-Services and Mobile Applications* (Kranz et al., 2014c).

[10]Parts of this chapter are based on a joint working paper with Florian Teschner and Christof Weinhardt which is currently under review.

[11]Parts of this chapter are based on joint work with Florian Teschner, Philipp Roüast, and Christof Weinhardt which has been presented at the *European Conference on Information Systems 2014* (Kranz et al., 2014a).

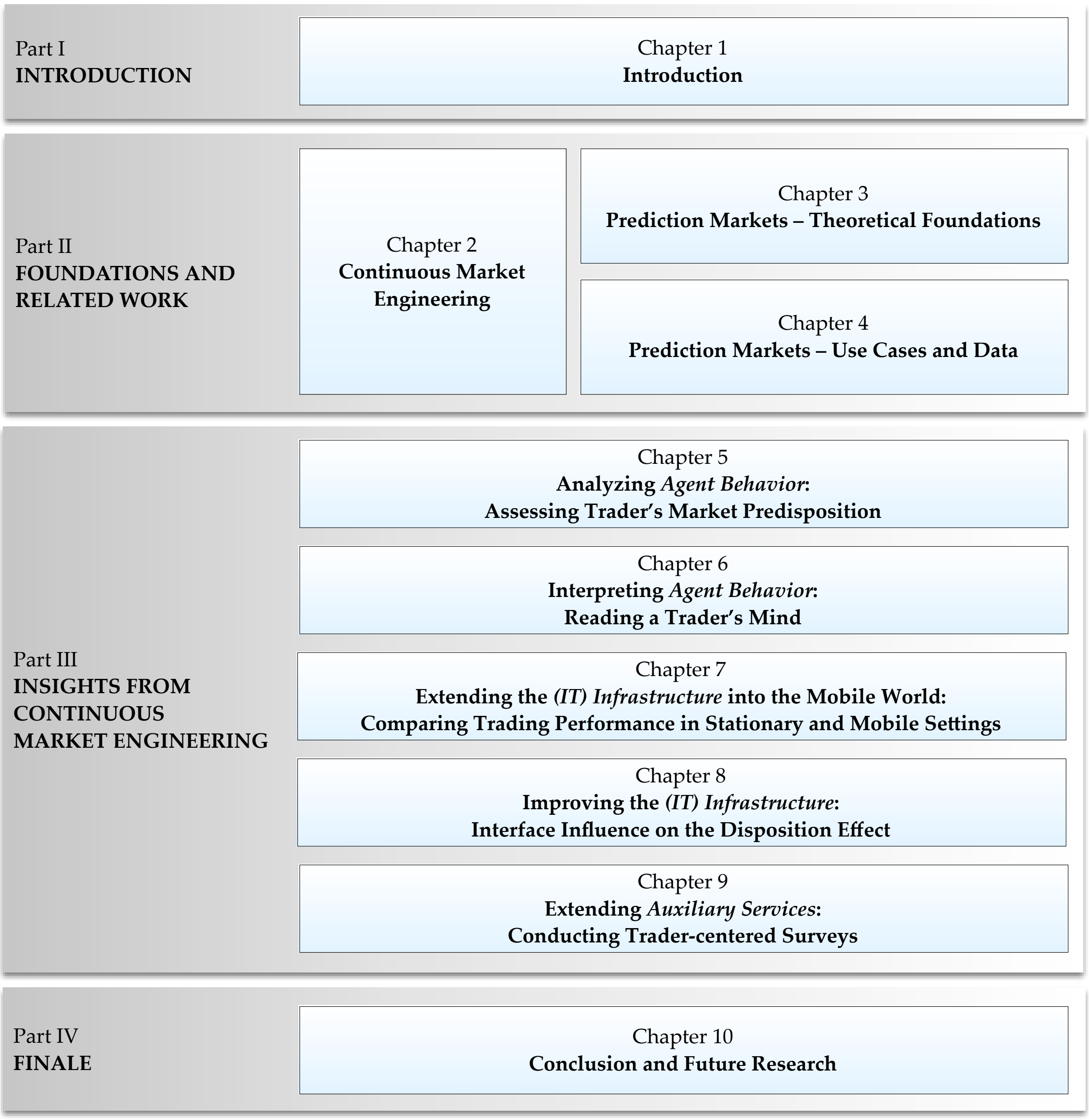


FIGURE 1.2: *Structure of this Thesis*

Part II

Foundations and Related Work

Chapter 2

Continuous Market Engineering

> "[...D]esign is important because markets don't always grow like weeds — some of them are hothouse orchids."
>
> ALVIN E. ROTH, 2002

2.1 Introduction

DESIGNING markets is a major challenge for market engineers, as hinted by the introductory quote. A meaningful market may seldom have a purpose of its own but often is a plain tool to achieve a certain goal, which in turn determines a market outcome. Thus, it seems fruitful to take a step back in order to assess what should be achieved, before prerequisites and goals can be translated to market characteristics. Furthermore, before a profound proposition on how to design 'good' markets can be suggested one has to fathom a market's typical lifecycle and ecosystem. Hence, this chapter starts by introducing a stylized market lifecycle followed by selected literature dealing with the design process of market systems. First, Roth's (2002) ideas on *Market Design* are presented, followed by *Market Engineering* (Weinhardt et al., 2003) and one of its derived forms, *Agile Market Engineering* (Block, 2010). These approaches may help in supporting a market engineer in creating markets to achieve desired goals. Finally, the significance of a continuous approach to the design and improvement of markets, called *Continuous Market Engineering*, is highlighted.

2.2 A Typical Market Lifecycle

Literature describing and developing models of lifecycles are around since many decades and were applied to many research domains; starting from the biological sciences, but also in social sciences (e. g., O'Rand and Krecker, 1990), in organizational sciences (e. g., Quinn and Cameron, 1983), object-oriented programming (e. g., Henderson-Sellers and Edwards, 1990), regional clusters (e. g., Fornahl and Menzel, 2007) or to describe typical phases in a product's life (e. g., Vernon, 1966; Wells and Gubar, 1966; Segerstrom et al., 1990). Typically, an electronic market comprises multiple components making up the market as such; i. e., the market software, computer systems, etc. Additionally, auxiliary services such as customer service processes or billing processes usually support the market system. Even though, an electronic market is not a product in the classical sense, it befells a lifecycle and can hence be matched to a product lifecycle fairly well.

The lifetime of a product is usually separated into phases which are traversed throughout a product's lifecycle. Figure 2.1 depicts the five phases of a product's lifecycle as described by Stark (2011), products – including services, and thus markets – run through in their existence: 'Imagine', 'Define', 'Realise', 'Use/Support', and 'Retire/Dispose'. He further groups those activities as follows. First, *Beginning-of-Life* (BOL) of a product comprises the three phases 'Imagine', 'Define', and 'Realise'. Second, *Middle-of-Life* (MOL) is described by 'Use/Support' of a product. Third, the phase 'Retire/Dispose' makes up the *End-of-Life* (EOL) of a product.

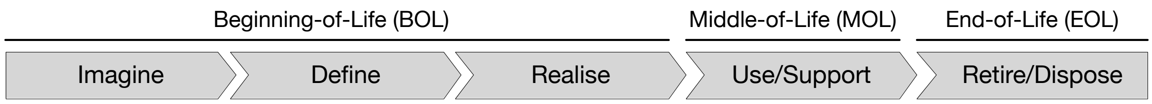


FIGURE 2.1: *Product Lifecycle (based on Stark, 2011)*

The *Stage-Gate System* by Cooper (1990), as depicted in Figure 2.2, is based on a slightly different lifecycle model. It separates the BOL in five stages, which are each connected with so-called gates, fulfilling quality assurance functions that will not be discussed here. Every new product starts with an idea, and enters the first stage 'Preliminary Assessment', followed by the stage 'Definition'. Afterwards, the product will enter the stage 'Development' and pass on to stage 'Validation'. The last stage is called 'Commercialization' and deals, inter alia, with production, sales, and related processes. Both Cooper's and Stark's phases of a product's BOL find their equivalence in the Market Engineering Process, as depicted in Figure 2.4b, that will be discussed in the next section. However, the MOL and EOL are described rather shortly with one phase each in the model of Stark (2011). As Cooper (1990) explicitly concentrates on the BOL, MOL is only broached, and EOL is not discussed at all.

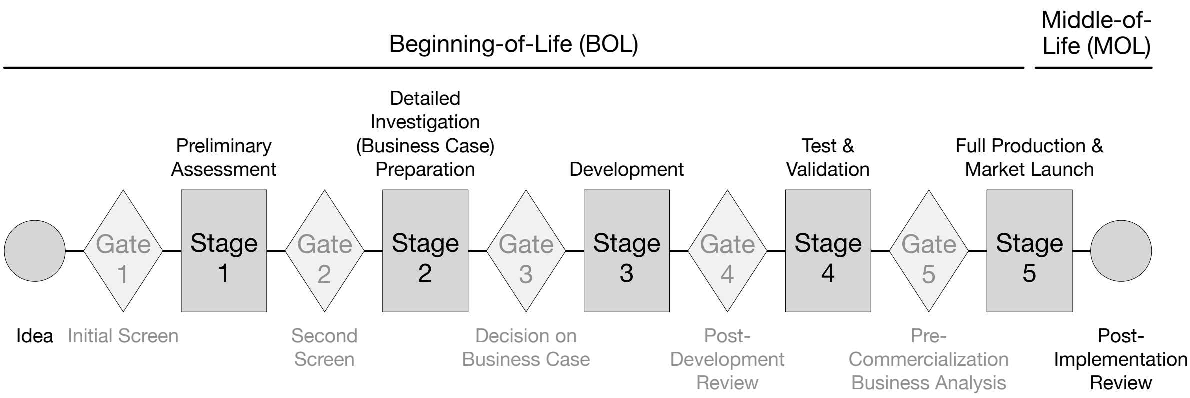


FIGURE 2.2: *Stage-Gate System (based on Cooper, 1990)*

A more detailed depiction for the MOL is the stylized market lifecycle as shown in Figure 2.3, which comprises four phases.[1] It describes, in a simplified way, the alteration of sales and profits of a market during its utilization. The caption 'sales' of the upper curve in Figure 2.3 is not meant literally, but represents the turnover generated with the market – the analogy to sales in a manufactured product context – rather than realized sales on the market. In the first phase, the 'Emerging phase', the yet unknown market is introduced to the intended audience. Tasks in that phase include the involvement of marketing methods in order to advertise the market. Nevertheless, the market is already fully operative and is used by a small, but steadily increasing number of participants.

Although, the market does already generate some sales, it generally yields negative profits. This changes in the following 'Growth' phase, where the market starts to return

[1] Even though called 'Market Lifecycle', this prototypal view can also be applied to a general product context and under certain assumptions also to a service context.

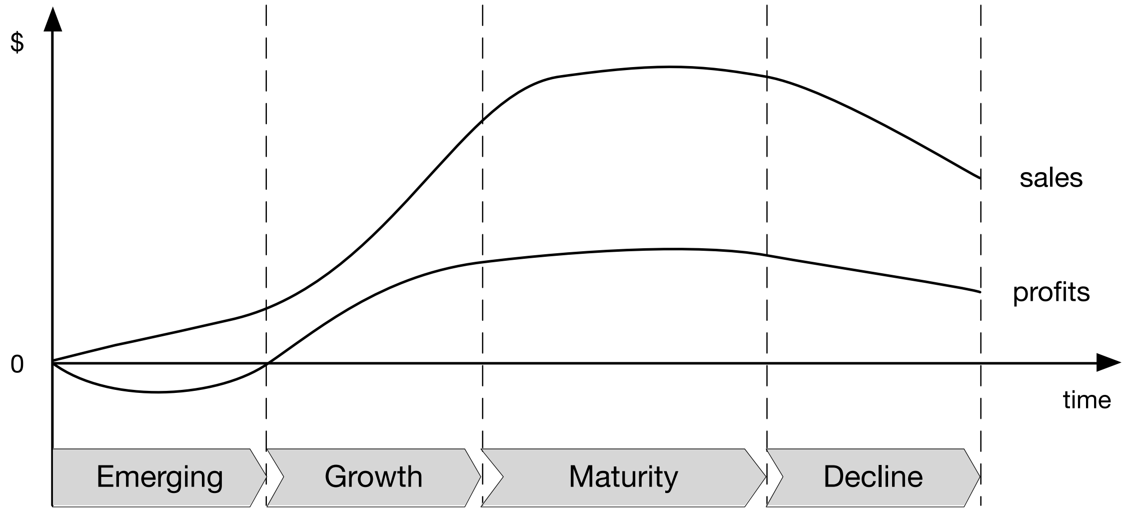


FIGURE 2.3: *Stylized Market Lifecycle*
(based on Weinhardt and Gimpel, 2007)

positive profits. Furthermore, the sales' growth rate increases until the sales reach an inflection point from which the phase of 'Maturity' starts. This is generally the phase in which most profits are yield and least resources need to be invested. The profits and sales tend to stay largely constant. Afterwards, at a certain point, the sales start to decrease, followed by diminishing profits; indicating the start of the 'Decline' phase. Finally, the market reaches its EOL and is discontinued. Summing up, the presented lifecycle models assume a linear, serial and thus non-repetitive phase model for products. They are well suited for a simplified descriptive reflection of product lifecycles and thus furthers the understanding of basic patterns in a product's lifetime. Nevertheless, they do not focus on the possibility to re-iterate certain phases in order to adapt a product in a dynamic environment.

2.3 Market Design

In his seminal paper, Roth (2002) discusses the changing expectation on economists to not only analyze markets, but to design them. Hereby, he implicitly calls for new methods, guidelines and frameworks for market designers. This very change from an analytical ex-post based perception to a formative ex-ante way of thinking implies the need for new approaches. Hence, research should not only concentrate on a conceptional, characteristics-based track, but extend efforts to investigate cause-effect relationships of market features to eventually enable an anticipatory design approach. Furthermore, he postulates two theses guiding the advent of *Market Design*. First, *[. . .] in the service of design, experimental and computational economics are natural complements to game theory"* (Roth, 2002). Second, he states that *[. . .] we need to foster a still unfamiliar kind of design literature in economics, whose focus will be different than traditional game theory and theoretical mechanism design"* (Roth, 2002). In 2008, Roth concretizes his earlier work, where he also postulates three prerequisites for markets to flourish: *thickness*, *congestion*, and *safe*ty. First, *thickness* must be provided by markets in the sense of attracting a sufficiently large share of a market's target group. Second, markets need to have a certain level of liquidity in order to overcome potential *congestion*. Third, markets must provide *safe* and simple trading opportunities to prevent participants from choosing any outside option. Finally, he corroborates his approach to market design on the basis of detailed examples of medical and academic labor markets, kidney exchanges, and school assignment problems. *"After a market has been designed, adopted, and implemented, it has a continuing life of its own"* (Roth, 2010). In the light of the latter quote, Roth (2010) revises his earlier publication (Roth, 2008) and reports developments and insights gained so far. Although, the presented markets still operate effectively, he nevertheless identifies unsolved problems that call for solutions. Milgrom (2011) sees market design closely linked to mechanism design, yet admitting that relevant questions of market design such as product definition, communicational concerns, incentives, and linkages among markets are not well studied in the context of mechanism design.[2] Examples of applied mechanism design in areas of market design include online advertisement markets (e. g., Edelman et al., 2007; Levin and Milgrom, 2010), electricity markets (e. g., Wilson, 2002), or radio spectrum auctions (e. g., McMillan, 1994).

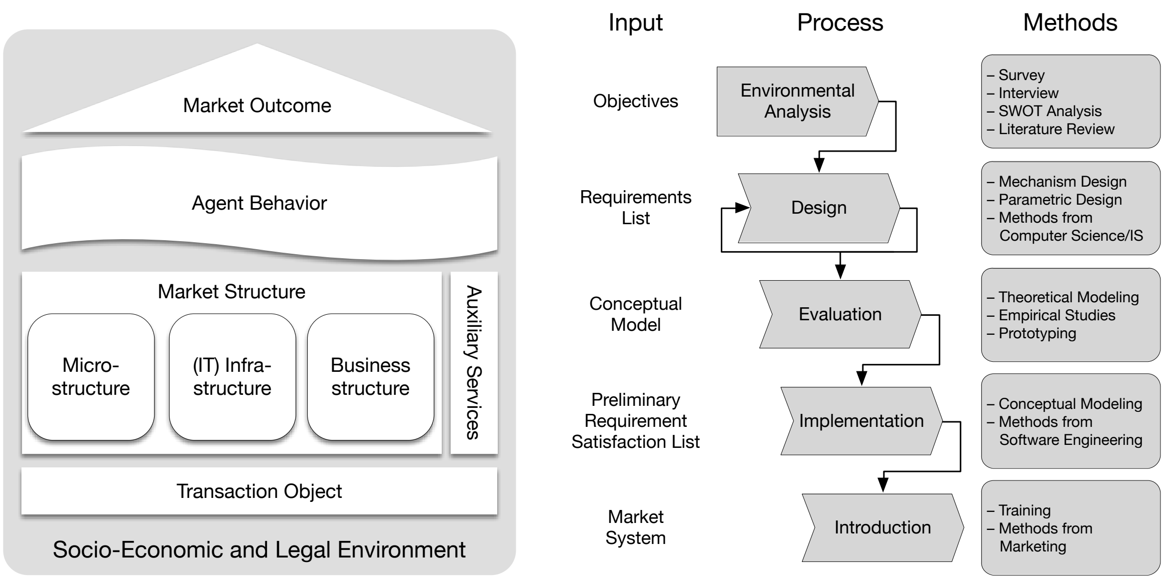
Summing up, *"[. . .] market design calls for an engineering approach"* (Roth, 2002) and stresses that details matter in relation to the market's mechanics and thus its purpose. Stated differently, it lacks a structured procedure model, capable of guiding a market designer in the creation process.

2.4 Market Engineering

Market Engineering as defined by Weinhardt and Gimpel is *"[. . .] the process of consciously setting up or restructuring a market in order to make it an effective and efficient means for*

[2] *"Market design is an engineering discipline linked to mechanism design."* (Milgrom, 2011)

carrying out exchange transactions" (Weinhardt and Gimpel, 2007).[3] It requires *"[...] an integrated, [holistic] view on markets, a multiplicity of methodologies, an interdisciplinary approach, and the understanding that details matter"* (Weinhardt and Gimpel, 2007). Market Engineering can be seen as an advancement to Market Design as described in the preceding section, although it is slightly different. In contrast, to Roth's (2002) focus on the application of market mechanisms, Market Engineering describes a systematic development process for markets (Market Engineering Process, cf. Figure 2.4b) besides a framework (Market Engineering Object, cf. Figure 2.4a) which supports market engineers in structuring the creative leeway. It hereby focuses on implementation details whilst considering environmental conditions and influences on agent behavior in order to reach the desired market outcome.



(A) *Market Engineering Object*

(B) *Market Engineering Process (selection)*

FIGURE 2.4: *Market Engineering Framework (based on Gimpel et al., 2008)*

The Market Engineering framework describes a *Market Engineering Object* (Figure 2.4a), consisting of multiple components, which gives an orientation for market engineers as well as a *Market Engineering Process* (Figure 2.4b) that describes the necessary steps in creating a market. All components of the Market Engineering Object are embedded within a *Socio-Economic and Legal Environment* which comprises applicable laws, social norms,

[3]Another definition can be found in Neumann (2007): *"Market [E]ngineering is the engineering design of all institutional rules of an electronic market"* (Neumann, 2007).

and others. The Socio-Economic and Legal Environment is considered as given and can usually not be modified by the market engineer. A market engineer's aim is to achieve a certain *Market Outcome* (e. g., activity, liquidity, allocation efficiency). This goal can be reached by designing the *Market Structure*, the *Transaction Object*, and the *Auxiliary Services*. *Market Structure* encompasses three interdependent components: *Microstructure*, *(IT) Infrastructure*, and *Business Structure*. *Microstructure* defines the market mechanism, like allocation and pricing rules. The *(IT) Infrastructure* comprises technical implementation details and interfaces to the market. Finally, trading fees as well as business and pricing model details belong to the *Business Structure*. Here, one has to be aware of the fact that the three aforementioned components, which make up the Market Structure, are strongly interdependent and thus cannot be designed independently of each other. *Agent Behavior* results from those components as well as from certain agent characteristics (cf. Chapter 5). In other words, it is not feasible for a market engineer to *directly* influence *Agent Behavior*. Therefore, he has to anticipate a certain *Agent Behavior* as result of his design decisions in order to achieve a desired *Market Outcome*.

The *Market Engineering Process*, based upon waterfall models as known in software engineering, leads a market engineer in creating a market. First, subject to the market objectives, the requirements are gathered by means of an *Environmental Analysis*. Second, based on the derived requirements list, the market *Design* is created. Afterwards, the *Evaluation* of the market is conducted on the basis of a conceptual model of a market system. Depending on the outcome of the evaluation step, the process re-iterates the design step or continues with the *Implementation*, where the market is created based on the preliminary list of requirements to be satisfied ('preliminary requirement satisfaction list'). Finally, *Introduction* of the market marks the last step of the Market Engineering Process. Weinhardt and Gimpel (2007) characterizes the Market Engineering Process as a *"basically sequential process"* that, nevertheless, allows iterations of process steps. Albeit acknowledging the necessity of iterations, the discussion on conditions and requirements thereof is not further stressed.

A juxtaposition of the main phases of the Market Engineering Process (Figure 2.4b) besides those of the discussed lifecycle models from Section 2.2 is depicted in Figure 2.5. As depicted, the Market Engineering Process supports the Market Engineer in each step throughout the Beginning-of-Life phase. Similar to the Stage-Gate System and the Product Lifecycle Phases introduced earlier, the Market Engineering Process does hardly concern the Middle-of-Life nor End-of-Life phases.

Market Engineering has successfully been applied to the creation process of markets in

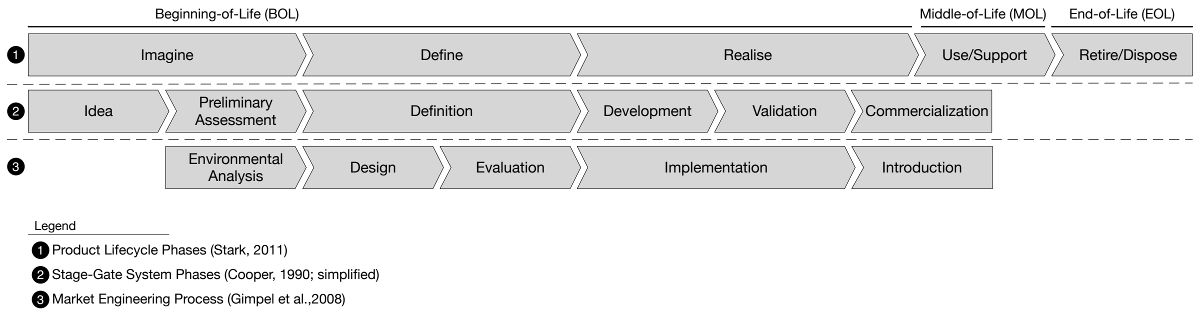


FIGURE 2.5: *Juxtaposition of Lifecycle Models in the Beginning-of-Life (based on Cooper, 1990; Weinhardt and Gimpel, 2007; Stark, 2011)*

various domains like (computing) grid markets (e. g., Schnizler et al., 2004, 2008), sport prediction markets (e. g., Luckner et al., 2005), economic indicators (e. g., Teschner et al., 2011), or environmental predictions (e. g., Stathel et al., 2009). It has been proven suitable in supporting market engineers in the creation process of markets. Nevertheless, albeit it allows for re-iterations, it lacks a focus on continuous monitoring and improvement of an existing market.

2.5 Agile Market Engineering

Block (2010) suggests *Agile Market Engineering* as a specialized advancement to Market Engineering. It contains a development process model including propositions for responsibilities[4] and software development methodologies as well as a collection of accompanying software tools. In its core, Agile Market Engineering is still based on aspects of the aforementioned Market Engineering framework, especially it makes use of the Market Engineering Object. Its main contribution is to describe an agile development process in detail, which is enriched with the experienced knowledge gained from practical market development projects. The *Agile Market Engineering Process* is depicted in Figure 2.6 and shall be understood as a "*[. . .] collection of best practices, experiences, and tools and furthermore shows one way to orchestrate them [. . .]*" (Block, 2010). It comprises three phases, namely the *Pre-Development Phase*, the *Development Phase*, and the *Operation Phase*, whereas each of these phases consists of three process steps.

[4]The proposed responsibilities/roles are business owner, market developer, market expert, market operator, change manager, and market participants. For further details see Block (2010, pp. 49)

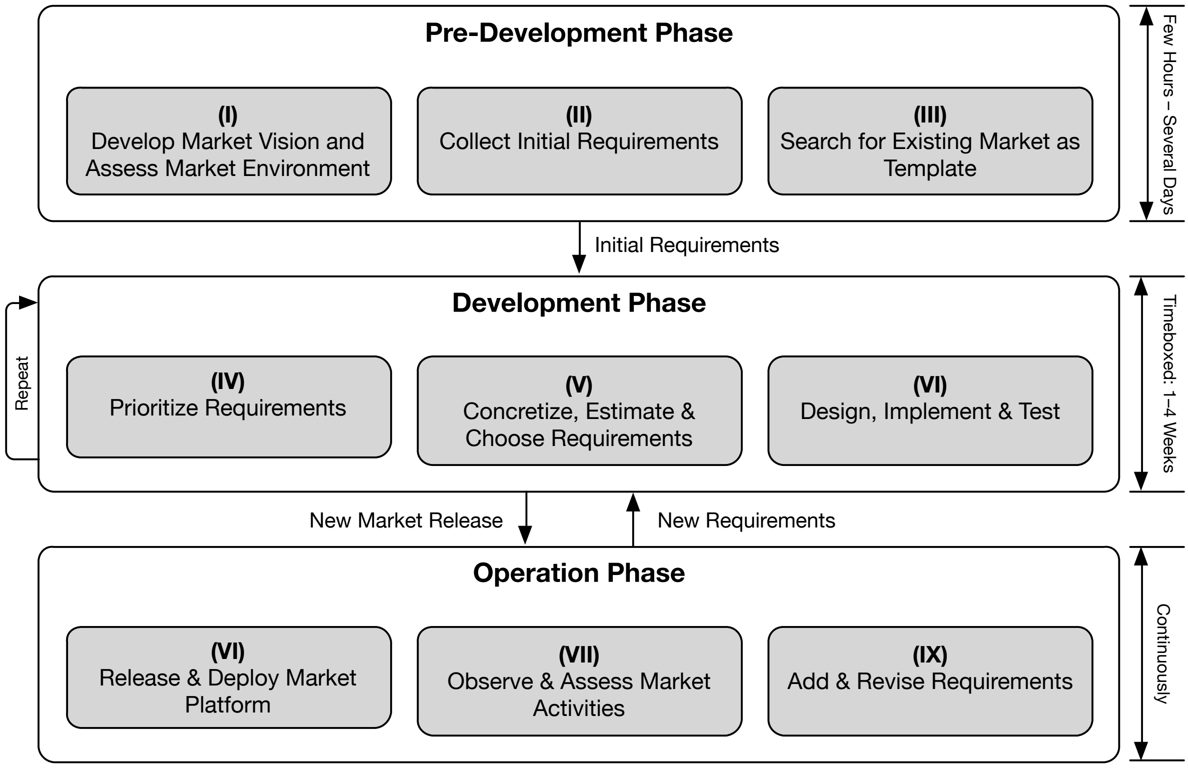


FIGURE 2.6: *Agile Market Engineering Process (based on Block, 2010)*

First, in the *Pre-Development* phase, the *Business Owner* develops a vision for the market and assesses the market's environment. For the first subtask, Block (2010) proposes to make use of creativity techniques like freewriting (Belanoff et al., 1991), brainstorming (Osborn, 1953), or mindmapping (Buzan and Buzan, 1993). Second, the *Business Owner*, assisted by the *Market Expert*, collects initial requirements for the upcoming market. He suggests to rely on user stories (Beck, 2000; Jeffries, 2001) for that purpose. Third, the *Business Owner*, accompanied by the *Market Engineer* as well as the *Market Developer* evaluate potentially existing market platforms by their similarity to the market's vision, which may act as a starting point for development. Therefore, Schönfeld and Block (2010) developed a market template repository. According to the Agile Market Engineering Process, it will take a few hours to several days to complete these steps. As result, a set of initial requirements is obtained, which initiates the *Development* phase. The steps of the development phase may be repeated numerous times. Each iteration is expected to last from one week to one month. First, the requirements are prioritized by the *Business Owner*. For this purpose, it is suggested to make use of prioritization schemes like MoSCoW (Clegg and Barker, 1994; Ash, 2007) or the one Selhorst (2006) developed on the basis of Kano et al. (1984). Second, the *Business Owner* along with the *Market Developer* finalizes the requirements and chooses a subset of user stories to implement in the following phase. Third, the *Market Developer* with the backing of the *Business Owner*, the *Market Expert*, and the *Change Manager* design, implement, and test the current set of user stories. Here, a test-driven development process (Beck, 2002) is proposed. This phase ends with a release of the market system. Afterwards, the *Operation* phase, which runs continuously, starts. First, the most recent release of the market system is deployed on the production system by the *Market Operator* and the *Market Developer*. Second, the *Market Operator*, the *Business Owner*, and the *Market Expert* observer and assess the market's activity. This aims to ensure proper operation of the market system in both the business and the technical dimension, which also includes monitoring for fraudulent activity for which approaches like the one described in Blume et al. (2010) are eligible. Third, the *Business Owner*, the *Market Developer*, the *Market Expert*, and the *Market Operator* revise the market's requirements based on their experiences gained in the latter step as well as on implicit (Kelly and Teevan, 2003) and explicit feedback. Those revised requirements will potentially trigger a new iteration of the *Development* phase, which in turn eventually leads to a new market release.

2.6 The Need for Continuous Market Engineering

In contrast to the simplified concept that markets are planned, implemented, run, and finally closed, reality often narrates a different tale. Also, the 'classical' lifecycle phases of emerging, growth, maturity, and decline are an abstraction which is too simplified for many areas of application. On the one hand, sometimes markets turn out to be far more successful during their operation than expected in the planning phase. Thus, their runtime is often prolonged which in turn raises the probability for changes in market's environment or objectives that will create a necessity for adjustments. The opposite can also be observed, and calls for an analogous approach of identifying and analyzing (design) weaknesses and conducting subsequent improvements. On the other hand, markets can be planned with an open end date from the beginning, knowing that their initial concept is rather a first approximation than a final design, and conscious of the need for further adjustments. Both situations described often require a regular or even continuous revisiting of the market's initial objectives and requirements as well as the way they are implemented. In these cases, the Market Engineering approach, although it explicitly allows the possibility to re-iterate process steps of the Market Engineering Process (i. e., re-design of markets), does not stress the continuous aspect satisfyingly with the relation to reality's needs. Especially the Market Engineering Process (cf. Figure 2.4b) lacks explicit advice for a market's Middle-of-Life phase (cf. Section 2.4) and thus for a continuous monitoring and improvement process. In particular, a structured process address changes is not contained in the Market Engineering framework.

As already quoted from Roth (2010)[5], after its development a market is not simply released into a static reality, but rather in a dynamic context in which the market might not necessarily be able to fulfill its purpose without proper readjustments. Weinhardt and Gimpel (2007) describe this circumstance as a *"fusion of design- and runtime"*[6] and stress the advantages of these cases by stating that *"[...] service operators can continuously experiment with subsets of their user groups [...] and the real-time feedback allows continuous improvement in the design of their online businesses"*, eventually leading to *"a competitive advantage over potential new entrants"* (Weinhardt and Gimpel, 2007). Roth (2010) further stresses the importance of continuously monitoring running markets in order to ensure that they fulfill their purpose: *"For those involved directly in the market, it is useful to continue to monitor it to make sure it is functioning well"* (Roth, 2010). Especially in the context of prediction markets, it is known that *"[m]arkets can fail and have been observed to produce anomalous behavior (Thaler, 1993; Thaler and Ziemba, 1988) thus understanding how to design prediction markets for successful deployment to minimize these failures is critical"* (McHugh and Jackson, 2012). In turn, the Agile Market Engineering Process contains this very step by (potentially) continuously reiterating the development and operation phases in turn. This resembles a potential connection between 'Introduction' and 'Environmental Analysis' in the Market Engineering Process.

This thesis suggests a *Continuous Market Engineering Process*, as depicted in Figure 2.7, based on many years of experience gained while operating the prediction markets 'Economic Indicator eXchange', 'Political Indicator eXchange', and 'Kurspiloten' (cf. Chapter 4). The process was derived by a combination of describing the interplay of operating, monitoring and refining taking place on the aforementioned prediction markets as well as from knowledge and insights gained throughout this time. It hereby attempts to structure and abstract the approaches conducted and the lessons learned so far. Similar to the Agile

[5] *"After a market has been designed, adopted, and implemented, it has a continuing life of its own."* (Roth, 2010)

[6] Taken from the following quote: "For some Internet market platforms, like eBay and Amazon, an interesting tendency can be observed: after the initial introduction of the electronic market platform, there is no clear cut distinction between design-time and runtime any more. This equals the fusion of design- and runtime of other Internet services like Google, for example" (Weinhardt and Gimpel, 2007).

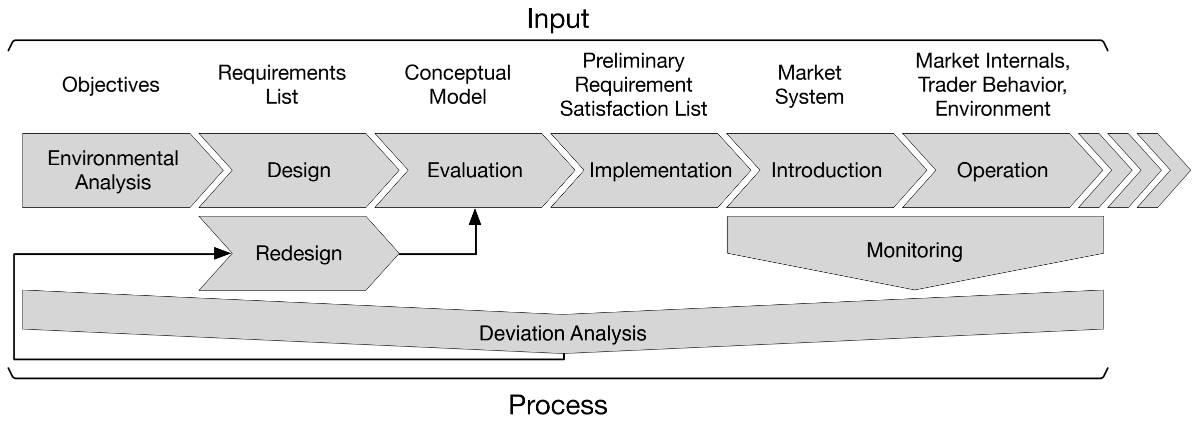


FIGURE 2.7: *Continuous Market Engineering Process*

Market Engineering Process, the Continuous Market Engineering Process can basically be viewed as an enhancement to the Market Engineering Process (Figure 2.4b) by aspects of continuous operation, monitoring, and refinement. In contrast to the latter, it does not end after the step 'Introduction', but moves on into the step 'Operation'. This step summarizes all activities of the market operator in running the market system (i. e., the market itself as well as auxiliary services supporting the market's operation). It is only interrupted by either (i) an introduction of a new release of the market system, which equals a restart of the step 'Operation', or by (ii) reaching the End-of-Life of the market, which results in a controlled suspension and closure of the market. The step 'Introduction' is accompanied by a 'Monitoring' step, which furthermore surveillances the 'Operation' of the market continuously. It describes all efforts undertaken by the market operator focusing on observing the market, i. e., the market's internals (e. g., technical operation of the market system, the market's metrics, etc.), the trader's behavior in the market, as well as environmental influences on the market (e. g., technological, legal and socio-economic changes). Monitoring the technical operation of the market system comprises monitoring the hardware and software the market system runs upon, as well as the market software system itself. Proper solutions may be built on top of common IT infrastructure monitoring solutions. By using market-specific key performance indicators, essential market metrics may be derived and conveniently monitored.[7] Semi-automatic monitoring tools a way to monitor trading behavior, like trading pattern recognition (e. g., Blume et al., 2010) or specialized logging facilities, which record key events. Environmental changes may be classified

[7]Generally, there are various possibilities to derive KPIs for a market. In case of prediction markets, a good starting point are the 'Market Quality Measures' as defined in the Prediction Market Quality Framework (Teschner, 2012, p.16).

as technological, legal, and societal changes. Examples for technological changes of the environment are accruing security issues, changes in trader's IT systems, or innovations rendering certain technologies obsolete. Legal changes, inter alia, comprise regulatory interventions forcing market operators to adapt their systems. Finally, socio-economic changes may affect traders' behavior and preferences, which in turn can result in different needs and demands on the market system. Each of the aforementioned monitoring goals are described rather abstract, as their particular design largely depends on the market's as well as the market provider's specific objectives. While a market is operating, usually the Continuous Market Engineering Process finds itself – besides in the 'Operation' step – in the 'Monitoring' step, continuously observing market parameters. There are basically two possibilities that will trigger a continuation of the Market Engineering Process. First, in case the 'Monitoring' step identifies a deviation from what is considered as normal operation, the step 'Deviation Analysis' is triggered, aiming to identify the underlying mechanics leading to this very change. Subsequently, it is decided how extensive the corrective actions have to be and thus triggering the 'Redesign' step, which needs to be initiated in order to refine the market. Second, changes of the market's objectives, resulting in the adaptation of subsequent input variables (namely, requirements, or conceptual changes; cf. 'Input' in Figure 2.7), can trigger the step 'Deviation Analysis' directly. For instance, a market operator decides to change its service portfolio (e. g., extending it by introducing a new class of products) will trigger the 'Deviation Analysis' directly. In turn, this process step will evaluate the available input parameters, eventually preparing the execution of the step 'Redesign'. Analogous to the Market Engineering Process, here the 'Redesign' step deals with the design of the market structure, as depicted in Figure 2.4a, followed by the 'Evaluation' step, which tests the market system with regards to functionality, acceptance and market outcome. For instance, reacting to a change in trader's activity by introducing additional transaction objects will result in a redesign of the market. Another example is the reaction to a security issue, which will result in a redesign of the market system's technical implementation. After conducting a 'Redesign' step, the process flow continues its way through the process model, reaching the step 'Introduction', which launches a new market release, and preliminary ends in the 'Operating' and 'Monitoring' steps.

2.7 Summary

This chapter presents different abstractions of lifecycles applicable for markets. Subsequently, the concept of Market Design is introduced before the Market Engineering framework, the related Market Engineering Process and an advancement, Agile Market Engineering, was discussed. Consistently, the importance of continuity in operating, monitoring, and re-designing markets in order to operate successful markets is stressed. Lastly, the *Continuous Market Engineering Process*, intended to guide market engineers in continuously operating and improving a market is presented. The derived approach of *Continuous Market Engineering* is applied subsequently to the work at hand.

Chapter 3

Prediction Markets – Theoretical Foundations

> "Prediction is very difficult, especially if it's about the future."
>
> NIELS H. D. BOHR

3.1 Introduction

PREDICTIONS about future events and scenarios which are reliable and trustworthy are often a prerequisite for economic as well as political decisions.[1] Common ways to gain insights about future developments are expert polls and mathematical prediction models using formal statistical methods such as regression analysis and ARIMA[2]. With the growth of the Internet, markets that are accessible via a web interface trading predictions about future events have emerged as a promising alternative forecasting tool. In these *Prediction Markets*[3], participants trade contracts with payoffs depending on the outcome

[1]Note, the words 'prediction' and 'forecast' are used interchangeable in this work.

[2]For a detailed description of the Autoregressive Integrated Moving Average (ARIMA) model see Box and Jenkins (1976).

[3]For an introduction to Prediction Markets containing a selection of successful applications see also Wolfers and Zitzewitz (2004).

of uncertain future events (see Section 3.2). Usually, participants submit their future expectations as orders like one does when trading stocks in financial markets. Prediction markets have been successfully applied to a wide range of topics such as sports betting (e. g., Borghesi, 2009) or political forecasts (e. g., Forsythe et al., 1992) as shown in Section 3.5.

Prediction markets derive their predictive power from a market design that provides incentives for traders to truthfully reveal their information and an algorithm which weights individual opinions. By aggregating dispersed information and expectations about future events into market prices, they generate a forecast and thus facilitate decision making. The price system plays an essential role to this end, whereto von Hayek (1945) wrote: *"The most significant fact about [the price system] is the economy of knowledge with which it operates, or how little the individual participants need to know in order to be able to take the right action."* In this context, the most important assumption of von Hayek (1945) states that gains from trades and equilibrium prices can be obtained even without common knowledge and perfect information. This holds especially for cases in which trader do not have (or reveal their) rational expectations, or whereby traders have only limited information of the markets state, or even if the number of market participants is rather 'small'. Smith (1982) examined three experimental market settings and found strong support for the so-called *Hayek Hypothesis*[4]. In a sense, the *Efficient-Market Hypothesis* (EMH), developed by Fama (1970), can be seen as a concretization of the Hayek Hypothesis. Simply speaking, von Hayek (1945) described the price system's capability to aggregate dispersed information into (equilibrium) prices even under suboptimal circumstances, whereas Fama (1970) hypothesizes that market prices already reflect all available information. Fama (1965) defines an 'efficient' market as a market *"[...] where prices at every point in time represent best estimates of intrinsic values."* According to the Efficient-Market Hypothesis, the fundamental value of a stock[5] in an efficient market is completely represented by the market price, which contains all relevant information of a stock and the market itself (Fama, 1970). In his 1970 paper, Fama breaks the EMH down into three subforms (*weak*-form efficiency, *semi-strong*-form efficiency, and *strong*-form efficiency) which he redefined slightly in 1991. In the weak form, past stock prices as well as other historic information do not completely determine future stock prices. Hence, past stock prices cannot predict future stock prices. Nevertheless, market participants cannot systematically exploit those inefficiencies to gain

[4]For a complementary study to Smith (1982) proving Hayek's hypothesis cf. Al-Ubaydli and Boettke (2011).

[5]The fundamental value of a stock, according to Harris (2003), is the value all traders would agree on, given complete information.

profits. Accordingly, stock prices follow a random walk model. In the semi-strong form, all publicly available information is – in addition to past price developments – incorporated into current prices. In the strong form, all relevant public and private information regarding the market are incorporated into market prices.

However, there are also preconditions a forecasting goal has to fulfill so that a prediction market can be used to generate a forecast for the particular event. Basically, a Prediction Market can be used to predict any future event that matches the following requirements: First, the event in question can be transformed into a number (i. e., a price). Second, as a forecast tries to describe the state of the given event at a specific date, it must be possible to determine the outcome of the event in question doubtlessly at the specified point in time.[6]

The remainder of this section, structured based on the *Prediction Market Framework* Teschner (2012) developed based upon Zhang et al. (2011), is as follows. First, the market microstructure is presented. Subsequently, the importance of a well-designed incentive scheme is stressed, followed by a discussion of ways to design the trading system as well as the market interface and their particular implications. Finally, exemplary use cases are presented.

3.2 Market Microstructure

Market Microstructure can be defined as *"[. . .] the study of the process and outcomes of exchanging assets under explicit trading rules"* (O'Hara, 1995, p. 15). First, the trading mechanism is presented before we turn towards the design of tradable contracts.

3.2.1 Trading Mechanism

Prediction markets neither rely on a specific trading mechanism nor are restricted to simple binary outcomes. The most basic trading mechanism for prediction markets is based on a *continuous double auction* (CDA) for a contract that represents the outcome of an upcoming event. Probably the most widely used market mechanism is a continuous double auction (Wolfers and Zitzewitz, 2004). Besides its low complexity and easy of implementation, a

[6]For details see paragraph on 'Contract Design' in Section 3.2.

CDA guarantees by design, that trading is a zero-sum game and hence is without financial risk to the operator.

If the number of participants and/or market liquidity is rather low or noisy (defined as 'thin markets'), a plain CDA is usually not an appropriate mechanism. Instead the use of market makers is beneficial under such circumstances.[7] Hereby, it is regardless of whether a human market maker or an automated market maker (i. e., an algorithm) is used. For instance, Hanson (2003) describes an automated market maker based on a market scoring rule to constantly provide trading opportunities. Stathel et al. (2009) discusses different market-maker approaches – amongst them the aforementioned one – and presents an own approach, which is also used in an intra-organizational prediction market presented in Stathel et al. (2010). This approach consists of an automated arbitrageur and an automated liquidity provider. The presented market (Stathel et al., 2010) comprises multiple stocks, each representing the probability of occurrence of a certain future scenario, including an outside option. According to the used payout function, the price of each product represents the probability of occurrence one to one (e. g., $ 12 equals 12 %). Hence, the sum of all stocks must relate to the sum of probabilities of occurrence of all future scenarios, which obviously is 100 %. Additionally, throughout the market runtime bundle trading was allowed.[8] By design, the market did not prevent traders from influencing the price of products so that the sum of open buy or sell orders lead to arbitrage opportunities. The automated arbitrageur described in Stathel et al. (2009) regularly checked for such arbitrage opportunities and leveled those by actively buying and selling as the circumstances require. The liquidity provider component of Stathel et al.'s (2009) market maker behaved similar to the one described by Boer et al. (2007). Generally, on the positive side, a market maker opens the possibility to provide rather small markets, but also has a downside. The market maker might cost some money to quote prices. In case of a human market maker, usually a salary and supplementary costs have to be paid, additionally both types of market maker might lose money when trading. Although, the maximum amount of money an automated market maker will cost in the worst-case scenario can often be calculated, a market engineer should be aware of that fact.[9] In marked contrast to most electronic trading systems, usually no trading fees are charged in Prediction Markets as this would have counter-productive effects on forecast accuracy, *"[. . .] because it inhibits*

[7]A market maker can either be an individual, an organization, or an algorithm, providing liquidity to the market by quoting both sell and buy prices.

[8]In the given setting, Bundle trading describes the market feature to sell one stock of each kind to the market system and receive $ 100 in return, or to buy one stock of each kind for a total price of $ 100 from the market system.

[9]For an in-depth analysis see Hanson (2007), Luckner (2008), and Chen and Pennock (2012).

rational participants from executing trades with an expected profit of less than the trading fee, thus, creating a bias" (Spann and Skiera, 2003, p. 1318).

3.2.2 Contract Design

There are multiple ways to design a contract in a prediction market, wherein the forecasting goal is the major determinant for that choice. The precondition to forecast any event with a prediction market is that the outcome can be quantified at a specific point in time. Wolfers and Zitzewitz (2004) describe three possible contract types. First, a *Winner-takes-all* contract that models events with a binary outcome, as for instance the chance that a certain candidate will be elected in a majority election. Second, an *Index* contract that is suitable to predict a numeric value representing a future event. Examples are the vote share of a certain party in a proportional representation or the number of unemployed persons at a given time (cf. Teschner et al., 2011). Third, a *Spread* contract which can be used to reveal the markets median expectation of a future event. In contrast to the aforementioned, it defines a fixed payoff and determine a particular spread via a market mechanism.

Any specific event and the corresponding tradable contract are connected via their *payout function*. The payout function is the key to market transparency as it enables participants to determine a priori what outcome equals which contract price, or as Antweiler (2012) expressed it: *"The ability of a trader to quickly convert and visualize the relationship between outcome and price is important."* Imagine a *Winner-takes-all* contract for a candidate 'A' in an upcoming election taking place with the method of majority decision. In this example, the payout function (see Equation 3.1) ensures one currency unit (1 CU) per stock in case candidate 'A' wins the election; else all stocks of that contract will be worthless.

$$p_A = \begin{cases} 1\,\text{CU}, & \text{Candidate 'A' wins the election} \\ 0\,\text{CU}, & \text{otherwise} \end{cases} \tag{3.1}$$

A more complex payout function for an *Index* contract might look like Equation 3.2.[10] Here, the inflation rate (I) is predicted and results in the price $p_{\text{Inflation}}$ as stated. The

[10] This example is based on Teschner et al. (2011).

relative difference to last years' value of I (i. e., $\frac{I_{t=0}-I_{t=-12}}{I_{t=-12}}$) multiplied by scaling factor α (here, 10) and added to 100 is paid out.

$$p_{\text{Inflation}} = 100\,\text{CU} + \alpha\,\text{CU} \times \frac{I_{t=0} - I_{t=-12}}{I_{t=-12}} \quad \text{, with } \alpha = 10 \text{ and } t \text{ in months} \tag{3.2}$$

Let us assume a rational risk-neutral market participant formed his expectation about the inflation development and derived with high certainty a relative change of 2.54 %. By using the payout function (Equation 3.2) the price per stock related to the given expectation can be calculated. In case the assumption was right, the resulting payout will be 125.40 CU. If the current market price of that contract will be below this value, the market participant should buy those stocks. In the opposite case, with market prices above 125.40 CU, the market participant should sell his stocks, since he expects them to be paid out for a lower price. In case of equality of expected payout value and market price, the market participant should be undecided whether to buy or to sell.[11]

Last, a *Spread* contract is explained in detail. It is again a zero-sum game for market operators and defines a fixed payoff which is paid when the event in question realizes. Here, as the amount available to pay out is determined by a fixed price both sides pay for a contract, market participants in a way negotiate the payout function itself. An example is depicted in Equation 3.3.

$$p_A^s = \begin{cases} 2\,\text{CU}, & s^* > s \quad \text{(Candidate 'A' wins more than } s\,\% \text{ of the vote)} \\ 0\,\text{CU}, & \text{otherwise} \end{cases} \tag{3.3}$$

Imagine a contract in form of an *even-money bet* (i. e., the winner's stake is doubled, while the loser receives nothing). Specifically, it shall be priced for \$ 1, paying \$ 2 if a election candidate 'A' wins a vote share s^* of more than s percent, \$ 0 otherwise. Hereby, traders submit bid and ask offers for the spread s, which in turn determines the potential payoff of that very contract.[12] For instance, a rational risk-neutral trader expecting candidate 'A' to win 30 % of all votes (i. e., $s^*_{expected} = 30\,\%$) would submit bids for that contract

[11]As the market participant's confidence will only in rare cases be 100 %, it is most likely, that he would sell his stocks in favor for cash than buying additional stocks that bear a risk of losing money. Furthermore, aspects like market activity and past price development – just to name a few – might play an important role in such a situation.

[12]Note, that s varies.

for less than 30 % and asks starting at 30 %, since $s^*_{expected} > s$ has to hold in order to double the stake. If bid and ask orders for s overlap, both traders pay the contract's price ($ 1). Furthermore, to conduct the final payout, the market system has to store the specific s of that transaction. After the event in question can be quantified, each transaction has to be paid out; i. e., it has to be determined, whether the (transaction-dependent) s is smaller than s^*. In turn, this decides which 'side' receives its doubled stake. Such a contract design follows that the bids and asks gathered in the order book will form the market's median expectation for s^*.

3.3 Incentives

Proper incentives are a crucial factor for a prediction market's success. Thus, an important part of a market engineer's work is to design and implement a suitable incentive system. There are numerous ways to classify incentives, like the very rough division into *intrinsic* (e. g., joy) and *extrinsic* (e. g., monetary reward) incentives or the finer-graded taxonomy of *financial*, *moral*, *coercive*, and *natural* incentives. In the context of prediction markets, it is common to make use of extrinsic (i. e., financial) incentives – also due to the fact, that creating moral, coercive or even natural incentives is a tough endeavor.

Regardless of whether a prediction market is run with real or play money, the market operator should provide monetary incentives. In case of play-money prediction markets, this can be achieved by supplying material prizes or vouchers. Besides that, incentives have to be communicated in a transparent way in order to be understood by participants and thus take full effect.

3.3.1 Currency

Every prediction market needs some kind of currency to enable trading. Using real money is by far the most straight-forward approach to fulfill that premise as participants as well as the market operators are used to it. On the downside, it can imply additional burden for market operators as they might – depending on their location – have to obey certain laws on lottery, gambling, or money laundering.[13] Additionally, using real money might

[13] Even though the *Socio-economic Environment* (cf. Chapter 2) can be assumed as given, it might be possible to circumvent those regulations. For instance, the market operator might call a single participation fee of € 10 for granting market access including 10 play-money units to participants.

discourage participants who are not willing to risk real-money although they want to share their estimations of future events. Furthermore, they additionally might be concerned to provide payment information. As research indicates, the predictive power of prediction markets – given a suitable incentive system – does not suffer, when using play money in combination with an appropriate incentive system (e. g., Servan-Schreiber et al., 2004; Christiansen, 2007; Slamka et al., 2008).

3.3.2 Tournament

Especially in case of play-money prediction markets, a ranking mechanism can be used to provide incentives for participants to reveal their true estimations. A public ranking can create a tournament-like setting for participants, that motivates a constructive and steady participation in the market. Furthermore, a ranking can be used to construct an additional incentive for play-money prediction markets. For instance, material prizes can be raffled amongst the best performing participants. In an insightful study, Luckner (2006) compared the influence of incentive schemes on prediction accuracy including rank-order tournaments. In contrast to naive expectation, he found that predictive accuracy was highest in the rank-order tournament treatment.

3.4 Technology

A prediction market is usually implemented as an online market, reachable from the Internet or in some cases as an intra-organizational market that is only accessible from an internal network (cf. Section 3.5 for an example). Besides aspects of usability, which are dealt with at the end of this section, questions of speed, scalability, and reliability are of interest from a trading system perspective.

3.4.1 Trading System

Multiple processes in a trading system are influenced by its speed. First, the number of orders that can be conducted per time unit gives an indication of throughput. Second, consistent with Riordan and Storkenmaier (2012), the time it takes from submitting an order until feedback about the order is available, after it is accepted and possibly executed from the trading system, indicates the latency of an order. *"Latency in electronic order-driven markets is determined entirely by the hardware and software (IT systems) used to match and report orders."* (Riordan and Storkenmaier, 2012, p. 417) Those measures gain importance when it comes to high frequency trading, or algorithmic trading in general. In case of solely human traders, the latency of a market system can be neglected, when it achieves a certain minimal quality that is far off modern stock exchanges.[14] Especially on prediction markets normally only human traders are involved and thus requirements for prediction market IT systems are usually moderate. The scalability of a trading system is a technical characteristic which is important for a market operator in order to assure a reactive market system, even under highly flexible demand patterns. Finally, the general reliability and safety of the trading system is a key prerequisite (cf. 'safe' in Section 1.1).

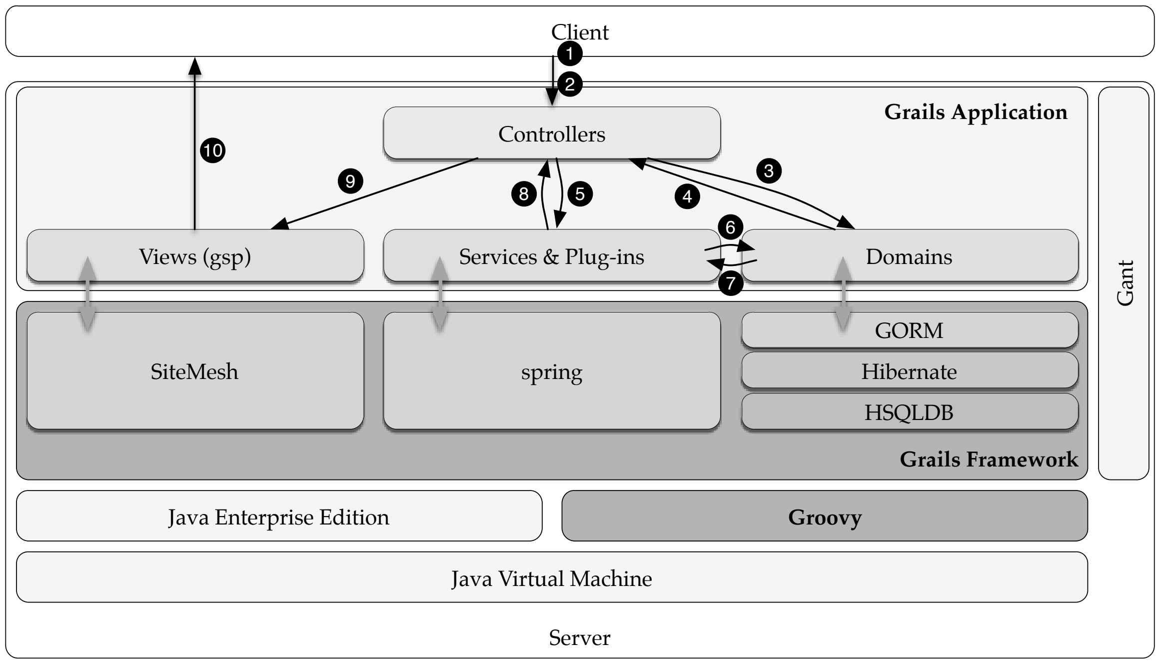


FIGURE 3.1: *GRAILS Framework*

An established framework meeting those requirements is *GRAILS*[15]. It builds upon established technologies like Java Platform Enterprise Edition[16], the dynamic programing

[14] For instance Deutsche Börse's Xetra system version 8.0 (released in 2007) reportedly already had an average latency of 10 milliseconds (according to Riordan and Storkenmaier (2012)), which is distinctly below the visual reaction time of an average person of 180-200 milliseconds (according to Kosinski (2013)).

[15] Accessible at the URL `https://grails.org/`.

[16] Accessible at the URL `http://www.oracle.com/technetwork/java/javaee/overview/index.html`.

language for the Java Virtual Machine Groovy[17], and the persistency framework Hibernate ORM[18] and others. The following example illustrates the process of requesting a web page from a GRAILS application (Figure 3.1): The client requests a specific page (i. e., URL) via a HTTP-Request (cf. Berners-Lee et al., 1996; Fielding et al., 1999) of the form http://Base-URL/Controller-Name/Action/Parameters (see label '1' in Figure 3.1). The application server routes that request to the indicated action of the specified controller including the specified parameters (see label '2'). Depending on the implementation of the requested action, the controller accesses or modifies necessary data from the data model (called 'Domain Objects' in grails terminology; see labels '3' and '4') backed by GORM. Thereby, the controller can make use of additional functionality offered by singleton instances available to all controllers ('Services' in grails terminology) or plug-ins[19] (see labels '5' and '8'). Services and plug-ins can themselves access and modify data from the model (see labels '6' and '7'). Next, the controller compiles the necessary data and passes it on to the corresponding view of the requested action (see label '9'). The view is responsible for generating a HTML file based on the data received (i. e., rendering). In that process, the view can make use of templates and additional programming logic to generate the HTML output, as indicted by the connection to the SiteMesh library in Figure 3.1. Finally, the application server replies the browser request with the generated HTML (see label '10'). A more specific example would be the submission of an order in one of those markets presented in Chapter 4. In that case, a trader would first login on the market system and would request the trading interface, analogous to the former example. When submitting an order, the client would send the necessary information as an HTTP request to the market system (see label '1'). The controller in charge for receiving orders would extract the order-related information from that HTTP request (see label '2'). Afterwards, it will generate a new domain object representing that order (see labels '3' and '4'). The domain object will be persisted by GORM according to the database settings configured in the respective application. Next, this very domain object is passed on a GRAILS service, that conducts the allocation, clearing and settlement of orders (see labels '6' and '8'). If the order can be matched to a standing order, the responsible service performs the necessary steps and saves the involved orders as well as writes and transaction log (see labels '5' and '7'). In case that the order cannot be matched, it will be saved as an open order and appear in the order book (see labels '6' and '7'). Finally, the controller passes the result on to the view (see label '9'), which renders a result screen, which is subsequently transfered

[17]Accessible at the URL `http://groovy.codehaus.org` .
[18]Accessible at the URL `http://hibernate.org` .
[19]For instance using the GRAILS mail plugin to send emails.

to the trader's client system (see label '10'). Both types of prediction markets discussed in the work at hand are implemented and operated with this very framework.

3.4.2 Interface

The market interface is a trader's major point of contact and interaction with a market and hence plays a pivotal role for a market's success. Furthermore, the interface can be used to 'nudge' (Thaler et al., 2010) traders into sensible behavior. Besides a form to submit orders, information about the market state are necessary for market participants to trade. Typical information presented to traders in a market interface are (i) name and description of the tradable contract, (ii) a chart showing past prices, (iii) an ordered list of unmatched buy and sell orders (order book), (iv) the last price at which two orders have been matched, (v) participant's holdings in that particular contract, (vi) participant's amount of cash available for trading, (vii) participant's open orders, (viii) additional information (e. g., price trends or order volumes). In case of an unexperienced market participant, this amount of information might be perceived as complex or even overwhelming. Thus, the question arises, whether every trader really needs every available piece of information and hence, if that complexity can be reduced. This question is primarily of importance when the negative effects of 'too much' information on individual's decision performance (i. e., *information overload*) are considered. In his 1984 paper, Jacoby examined information overload on customers and showed that consumers can be overloaded, albeit he was convinced that information overload will not occur. In a reply to Jacoby (1984), Malhotra (1984) was able to show empirically that information overload does occur in real-life settings, and thus that the simple relation 'more information is always better' does not hold. A more positive approach on the interdependencies of information and decision quality is the theory of *cognitive fit*. This theory examines the positive effects 'right' and 'appropriate' information has on decision quality. According to Vessey (1991), that a fit among problem representation and problem solving task leads to better and faster decision making. In a further study, Kleinmuntz and Schkade (1993) confirmed Vessey's (1991) main statement. A related school of thought is the *resource matching* theory (cf. Anand and Sternthal, 1989; Tan et al., 2010), which tackles the relationship of mental resources demanded for a certain task and mental resources available.

A major challenge of interface design is to present the right amount of information necessary for a market, suited to individual capabilities. As markets might also be used in domains where individuals do not expect them or find their application unnatural, one could alternatively take a step back and ask, whether it is necessary to confront individuals with a market at all. For instance, consider a fully dynamic pricing mechanism for electricity, in which consumers are unwilling to trade for every bit of electricity, but rather would use a rule based system through which they might set subjectively acceptable price bounds. Such a system would in a way encapsulate the market and thus hide complexity from the user. Against this background, Seuken et al. (2010) proposed the concept of *hidden markets*. In his work, he describes two sub-forms of hidden markets. First, *weakly hidden markets* attempt to *"find the right trade-off between hiding or reducing some of the market complexities while maximizing economic efficiency attained in equilibrium"* (Seuken et al., 2010, p. 1662). Second, *strongly hidden markets* completely hide some semantic aspects of the market. Even though hidden markets' main purpose was to increase user acceptance in domains where markets do not play a pivotal role, that concept is especially apt to be used in order to simplify a prediction market's interface as shown by Teschner and Weinhardt (2012).

Besides the trade-oriented market interface, a portfolio overview, a trading history and a profile page might supplement the interface. Furthermore, a prediction market usually contains additional static web content like game instructions and legal information.

3.5 Applications

Prediction markets have been applied to numerous domains since their appearance. This section presents a selection of exemplary applications of prediction markets, with the intention to provide an overview of interesting real-world use cases.

Political Stock Markets

Political Stock Markets (PSM) are one instance of prediction markets. They share their main objective, namely aggregating information from its participants in order to create efficient forecasts for uncertain future events. In this case, these uncertain future events are of political nature, i.e. elections, nominations for elections or policies.

Using prediction markets for political forecasts offers many advantages compared to traditional polls or expert surveys. As PSMs are usually open round the clock, participants can trade whenever they like and therefore react to news promptly (Snowberg et al., 2007). Since prediction market prices are updated immediately when participants incorporate their expectation by trading, PSMs provide continuously and timely updated forecasts. Usually, their market interface is interactive and the setting gamified, in marked contrast to most surveys, and thus providing further incentives for participation. Most surveys rely on random samples for validity and accuracy. In prediction markets, in comparison, those with the best information are the best participants – the very individuals who are most likely to self-select into the market. Additionally, as successful participants accumulate their profits they gain forecasting weight over time compared to less successful participants. With surveys, this process of self-selection would introduce a sampling bias, but with markets, the incentive system forces low performers out of the market in the long run. Turning to the disadvantage of markets over surveys, one has to mention the higher complexity burdening participants (Graefe et al., 2010). First, participants have to understand the trading mechanism. Second they have to understand how events are related to contracts. The process of understanding the 'task' (i. e., filling in a questionnaire or submitting an order) is more structured and better researched for surveys than markets. The forecast performance of prediction markets in general is still in debate. On the positive side, they have proven repeatedly to be very potent information aggregation mechanisms (e. g., Berg et al., 2008; Ledyard et al., 2009; Bennouri et al., 2011). Although, other evidence suggests that the relative performance advantage of markets may be small compared to surveys or polls (e. g., Goel et al., 2010; Erikson and Wlezien, 2008; Rothschild, 2009). Compared to eliciting expert opinions, prediction markets eliminate the effort of identifying experts and motivate their participation. In many cases they allow anonymous participation, which may increase the likelihood of nonconformists to participate and reveal information while they do not need to deal with conflicting opinions.

A famous example of PSM are the *Iowa Electronic Markets* (IEM)[20] operated by the University of Iowa launched in 1988 (cf. Forsythe et al., 1992). The question of PSMs' performance compared to polls has sparked some attention in the last years. Berg, Forsythe, Nelson, and Rietz (2008) analyze the results of more than ten years' worth of PSM predictions on the IEM against corresponding polls and conclude that market results outperformed the polls in most cases. Similarly, Berlemann and Schmidt (2001) find that – though by a less broad margin – European PSMs significantly outperformed respective polls as well. There has been some doubt with respect to the naive manner polls were used in their comparisons, i.e. Erikson and Wlezien (2008) argue that polls needed to be properly adjusted before comparison, but as Rothschild (2009) points out, fairly adjusting both PSM and poll results yields PSM as the overall more accurate predictor.

[20] Accessible at the URL `http://tippie.uiowa.edu/iem/` .

Economic Predictions

It has been known for long that the accuracy of classical economic forecasts are often disappointing (McNees, 1992). Even worse, inter alia, Heilemann and Stekler (2012) found that hardly any improvement in economic forecast accuracy could be observed in the last decades. With that in mind, Teschner et al. (2011) and Teschner and Weinhardt (2014) described a prediction market to forecast economic indicators called *Economic Indicator eXchange* (EIX). The predictive accuracy of selected macroeconomic indicators are promising and sometimes outperforming established forecasting methods (Teschner and Weinhardt, 2014). For a detailed description of the EIX see Chapter 4.3.

Sports Betting

One example of prediction markets used for sports betting is *Tradesports.com, Inc.*[21]. On that platform, numerous so-called contests for different forms of sport[22] are listed. A series of stocks is associated with each contest. Usually, the stocks are *Index* contracts (cf. Section 3.2). Tradesports.com was the object of a couple of prediction market studies targeting inter alia market efficiency, transaction costs, intra-game price movements, and disposition effect (e. g., O'Connor and Zhou, 2008; Borghesi, 2009; Hartzmark and Solomon, 2012; Borghesi, 2013). Sports betting prediction markets are similar to traditional odds-based sports betting platforms like Betfair[23] and others. In contrast to the aforementioned type of sports betting, the prediction market approach can deliver a more intuitive representation of current expectations, since the prices are usually easier to convert to probabilities. Furthermore, the rules platform providers apply for setting their odds commonly differ between sports betting providers as well as from the way market prices are derived in a prediction market.

Intra-Organizational Markets

Contrary to the ongoing development in the personal sphere of (self-)disclosing more and more information and thus eroding privacy[24], for companies information is often a valuable asset that has to be actively secured and kept undisclosed. Prediction markets are

[21] Accessible at the URL `http://www.tradesports.com/`.

[22] At the time of writing the following forms of sport are listed: American Football (NFL and NCAA FB), Baseball (MLB), Basketball (NBA), and Golf.

[23] Accessible at the URL `http://www.betfair.com/`.

[24] Remember Scott McNealy's famous 1999 quote (*"You have zero privacy anyway. — Get over it."*) or Marc E. Zuckerberg's 2010 claim that privacy is no longer a social norm (*"People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people. That social norm is just something that has evolved over time."*).

also suitable to be used in such settings that demand discretion. For instance, prediction markets can be used in a company to forecast whether an internal development project will be successful, on time or within budget. Another example is an intra-company prediction market used to assess and evaluate ideas and innovations (cf. Stathel et al., 2010). In both examples it is crucial to restrict access to the prediction market platform accordingly by technical provisions.

3.6 Summary

As has been shown in this chapter, prediction markets are not the ultimate off-the-shelf solution for forecasting challenges, due to their flexibility. However, properly designed prediction markets have shown to be successful in a variety of applications in the last decades (e. g., Forsythe et al., 1992; Zitzewitz, 2006; Huber et al., 2008; Berg and Rietz, 2010; Teschner et al., 2011). They have a long track of successful field applications, e.g., in political elections (e. g., Berg et al., 2008), sport events (e. g., Luckner and Weinhardt, 2008) , finance (e. g., Bennouri et al., 2011), innovation assessment (e. g., Stathel et al., 2010), and predicting market development (e. g., Spann and Skiera, 2003).[25]

The roots of their predictive power are twofold; prediction markets can provide incentives for traders to truthfully disclose their information and an algorithm by which to weight opinions. They facilitate and support decision making through aggregating expectations about forthcoming events (cf. Berg and Rietz, 2003; Hahn and Tetlock, 2005; Hanson, 1999).

[25] See Wolfers and Zitzewitz (2004) for a comprehensive review.

Chapter 4

Prediction Markets – Use Cases and Data

> "Prediction markets are remarkably accurate information aggregation mechanisms."
>
> STEVEN GJERSTAD, 2005

4.1 Introduction

THE studies presented in Chapters 5 – 9 are conducted on two prediction markets subsequently described here. First, the prediction market *Kurspiloten* is introduced, which is used in the studies detailed in Chapters 5, 7 & 8. Second, the Economic Indicator eXchange (EIX) and its submarket, the Political Indicator eXchange (PIX), which are used in the studies in Chapters 6 & 9 are introduced. Third, an overview is given about the runtimes of the different markets, the studies conducted, and the products tradable in different versions of the markets.

4.2 Kurspiloten

The *Kurspiloten*[1] market was the yearly stock-market game of the leading German business newspaper *Handelsblatt* in 2011 (cf. Table 4.2). It is a prediction market for selected stock

[1] engl. 'Quotation Pilots'

indices and commodities, developed at the Karlsruhe Institute of Technology (KIT) by the Institute of Information Systems and Marketing. The market is designed as a continuous double auction without a market maker, operating from 2011-09-02 until 2011-11-25. The cooperation with Handelsblatt helped to reach out to a broad and well-informed audience interested in financial markets and economic developments.

4.2.1 Market Design

The work presented in Chapters 5, 7 & 8 is conducted on a prediction market called *Kurspiloten*. It is a web-based prediction market designed to forecast the stock exchange value of selected stock indices and commodities on a weekly basis. Participants registered free of charge and traded with play money.[2] Therefore, they could not lose any real money. Prizes worth over € 70,000 were drawn among well-performing participants to incentivize them to reveal their true beliefs (cf. Section 3.3).

Like in financial markets, Kurspiloten is set up as a continuous double auction with one stock representing the final (real-world) price of one of the twelve predicted stocks at a given time. Six stock indices, three commodities, one commodity index, one future contract and one exchange rate can be traded (Table 4.1). The tradable contracts represent their underlying stock one-on-one. Participants are expected to buy if they think that current Kurspiloten prices underestimate their estimation of the final value of the underlying stock market index or commodity and sell if they think prices overestimate the final value. By trading their price expectations of twelve selected stock indices and commodities on a weekly basis, participants are able to share their private information with others.

Although Kurspiloten uses play money, participants are provided incentives to behave similar to a real-money market. Prizes worth over € 70,000 were drawn to well performing traders in order to provide incentives for truly reveal information. As the amount of play money was not extensible by some analogy of a deposit, participants had an incentive to economize their play-money budget (cf. Section 3.3). Hence, the dominant strategy of participants is to buy undervalued stocks and sell overvalued stocks. Furthermore, they should realize gains as well as losses in order to increase their buying power. The traders' total assets (i. e., total amount of money and stocks at market prices a trader owns) is used as a performance measure. Weekly prices worth around € 1,500 are awarded according to

[2]Due to legal restrictions the market had to rely on play money; nonetheless '€' was used as currency name. To avoid confusion 'P€', 'EIX€', or 'PIX€' are used in this thesis as currency sign for play money and '€' for real money.

TABLE 4.1: *Tradable Stocks on Kurspiloten Market*

Stock	ISIN	Underlying (currency, unit)
DAX	DE0008469008	30 major German companies (€, Index)
MDAX	DE0008467416	50 major German companies[a] (€, Index)
TecDAX	DE0007203275	30 largest German technology companies (€, Index)
EuroStoxx 50	EU0009658145	50 Eurozone companies (€, Index)
Dow Jones Industrial Average	US2605661048	30 major US companies ($, Index)
Nikkei 225	XC0009692440	Tokyo Stock Exchange (¥, Index)
EUR/USD	EU0009652759	EUR-USD exchange rate ($, €)
Euro-Bund Future	DE0009652644	Future contract on German national loan (€, €)
Gold	XC0009655157	Gold (€, Ounce)
Silber	XC0009653103	Silver ($, Ounce)
Brent Crude Oil	XC0009677409	Brent-Oil ($, Barrel)
Rogers International Commodity Index	NL0000424505	38 commodities from 13 international exchanges (€, Index)

Notes: In *Kurspiloten* market all stocks are traded in P€, regardless of the currency of their underlying (e. g., Nikkei 225 at ¥13,045 will have a payout value of P€ 13,045); [a]Excluding DAX and TecDAX

the ranking of the participants' assets at the end of each week. The main prize worth over € 40,000 is given to the most successful trader according to his overall assets, i. e., – since all stocks are paid out – the total amount of money one owns at the end of the game.

Upon registration each participant receives an initial endowment of P€ 100,000 and 1,000 stocks of each tradable asset. The trading period for all stocks is seven days. Each Friday at 5:30 pm, the market is closed for trading. To attenuate endgame effects the market is closed for trading five minutes prior to the payout. Afterwards all 12 products (Table 4.1) are paid out according to the stock exchange prices at 5:35 pm. All participants receive their new endowment consisting of 1,000 stocks each for the next seven-day trading period.[3] Finally the market is reopened for trading. As the experiment ran for twelve

[3]Due to a bad money/stocks-ratio a second account was introduced for each user called "Geldspeicher" (engl. Money Bin). Starting on 2011-10-23 all money exceeding P€ 10,000,000 is booked to the Money Bin in the weekly payout procedure.

weeks, 144 payouts were executed. Any order submitted for a paid out product can be rated ex post as 'informed' or 'uninformed' in relation to the payout price. For instance, the stock 'DAX 07.10.2011' was tradable from 2011-09-30 until 2011-10-07 at 5:30 pm and represents the (real-world) price of DAX on 2011-10-07 at 5:35 pm (GMT+1), which is 5,673.08. Imagine (a) a buy order for this stock with a limit price of P€ 5,715 and (b) a buy order for this stock with a limit price of P€ 5,660. Order 'a' is an uninformed order, since its limit price is higher than the payout price (i. e., the final value of the underlying stock) and will therefore most likely result in a loss. In contrast, order 'b' can be regarded as an informed order, since its limit price is below the payout price and thus its execution will result in a gain of P€ 13.08 per stock when it is paid out.

Registration for Kurspiloten was free of charge and open for anyone. In the registration process participants only had to enter a valid email address and a username. Participants could register up to three days before market opening, or at a later time. Due to the repeated endowments of stocks for following trading periods a trader receives after each payout, participants who registered after market opening would be disadvantaged. In order to give those traders a chance to catch up with the competitors, their initial endowment is adjusted to the account balance a hypothetical user who registered on the first day would have.[4] If a user registers after market opening, he receives the amount of money a hypothetical user who registered on the first day would currently own (i. e., all past endowments multiplied with their corresponding payout values) including the initial portfolio for the current week.

Proper incentives are set by using a public ranking list containing the usernames based on the traders' absolute assets. This ranking's primary use is to award the € 70,000 in prizes. The ranking was accessible for all traders throughout the market's runtime. Thus, the second incentive is a trading-performance dependent social comparison. In order to inform participants about the market rules, general instructions explaining the basic market rules and conditions of participation are provided. The instructions were neither individualized in any way nor adapted to the specific treatment a participant might be part of.

[4]This is achieved by creating a dummy user account on the first day of the market which receives the same endowments as each user and is paid out in the same way a normal user is.

4.2.2 Market Interfaces

Kurspiloten Web Interface

The trading interface of the Kurspiloten web interface is displayed in Figure 4.1. Participants have convenient access to their portfolio and account information (W1), market information (W2) such as the last trading day, the order book with five visible levels of depth (W3), and an up-to-date financial news stream (W4) provided by the Handelsblatt. As additional information, a chart of the Kurspiloten prices is displayed in comparison to the stock's price development (W5). Finally, a short reminder about the rules of the game is displayed (E).

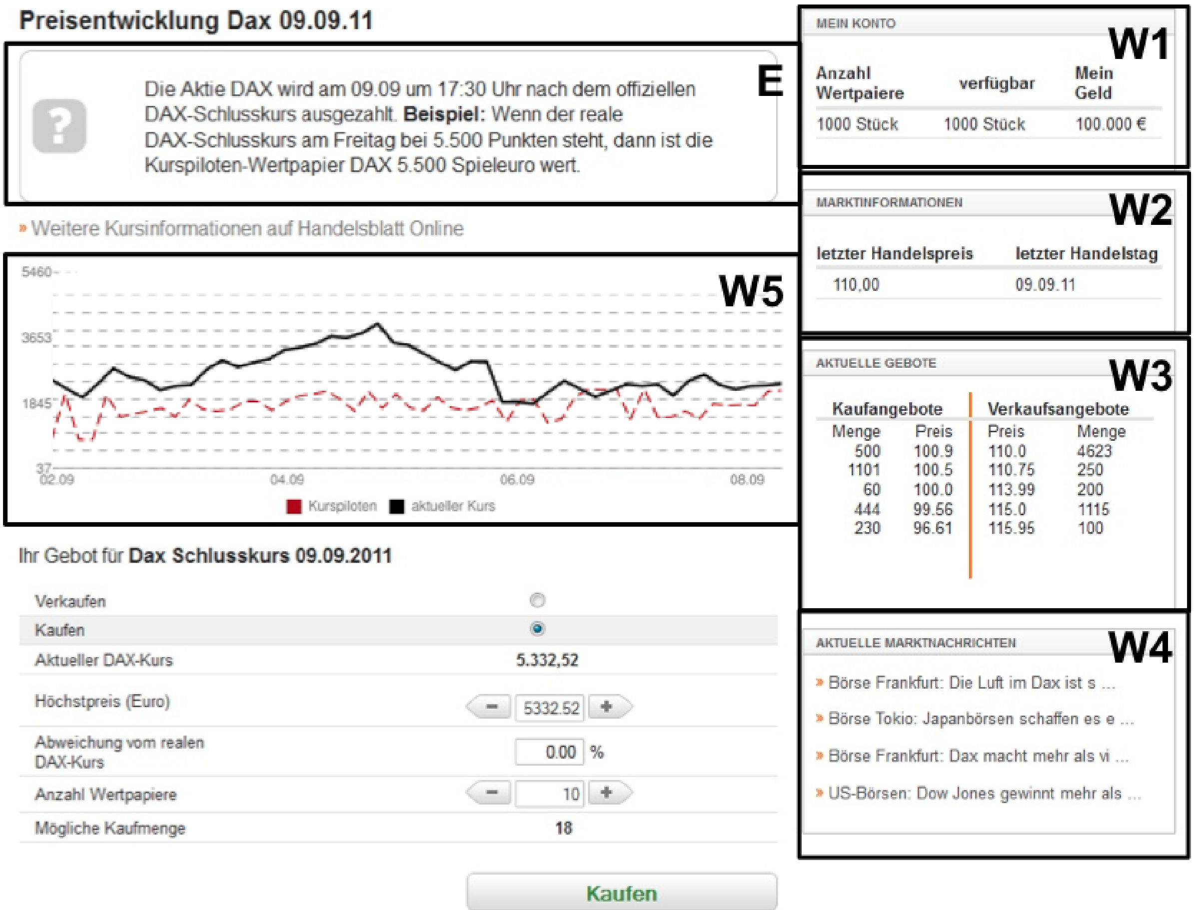
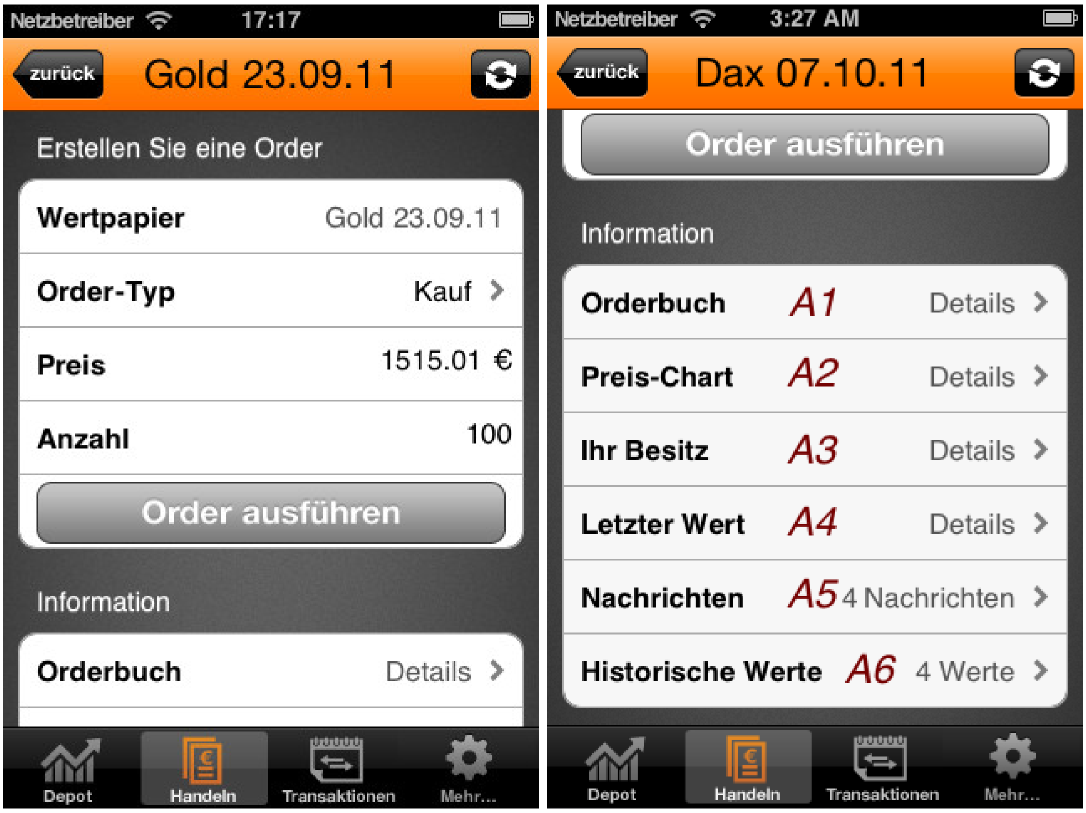


FIGURE 4.1: *Kurspiloten Web Trading Screen*

Kurspiloten Mobile Interface

The *Kurspiloten App* (KAPP) is a mobile client for the Kurspiloten Market designed for iOS version 4 and above of Apple's iPhone. KAPP's design is a compromise of two goals: First, KAPP is intended to be easy to use for all users. Therefore, the design followed Apple's iOS Human Interface Guidelines[5] to look and feel native on the iOS platform. Second, existing Kurspiloten-users should be able to use the App with minimal learning-effort. Therefore, KAPP uses the same wording, trading workflow, and – as far as possible – information elements as the Kurspiloten's web-interface. Due to the limited screen size of the iPhone platform, it is not reasonable to let KAPP look exactly like Kurspiloten's web-interface. Nevertheless, KAPP's frontend tries to be close to Kurspiloten's web interface inter alia by sharing the same menu structure and nomenclature.



(A) *upper part* (B) *lower part*

FIGURE 4.2: *Kurspiloten Mobile Trading Screen*

KAPP contains all core features of the Kurspiloten Market (Figure 4.2), i. e., submission

[5]Accessible at the URL `https://developer.apple.com/library/ios/documentation/userexperience/conceptual/mobilehig/`.

of orders, cancellation of orders, examination of own holdings in the portfolio and the available amount of money. Analogous to the Kurspiloten web interface, the KAPP trading screen offers access to six dedicated information screens (Figure 4.2b). Most of them are congruent to the web interface's information entities, namely the order book with five visible levels of depth (A1, cf. W3 in Figure 4.1), a chart containing the stock's real price development in comparison to the Kurspiloten price (A2, W5), the user's own holdings of the stock (A3, W1), the last value and closing time of the contract (A4, W2), and an interface to the same news stream as in W4 (A5). Additionally, participants have access to previous (real) stock values (A6). KAPP lacks some secondary features of Kurspiloten, namely the online-help and ranking list. KAPP was submitted to the App Store and could be downloaded at no charge while Kurspiloten was operational. In order to gain a broad user-base, an 'advertisement' – consisting of link to KAPP and a short description text – was put on the homepage of the Kurspiloten website.

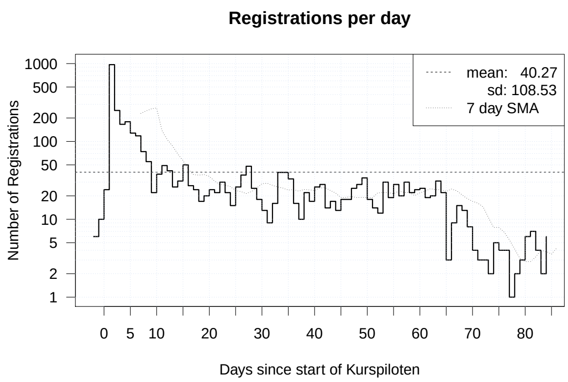
To allow research about users' information usage prior to submitting an order, the consumption of the six information screens are logged in KAPP (information entities A1-A6, Figure 4.2). Kurspiloten and KAPP both track the time a user needs to create and submit an order. As KAPP uses separate screens to display stock related information, it also records the user's consumption time of these information units on a per item basis.

4.2.3 Dataset and Descriptive Statistics

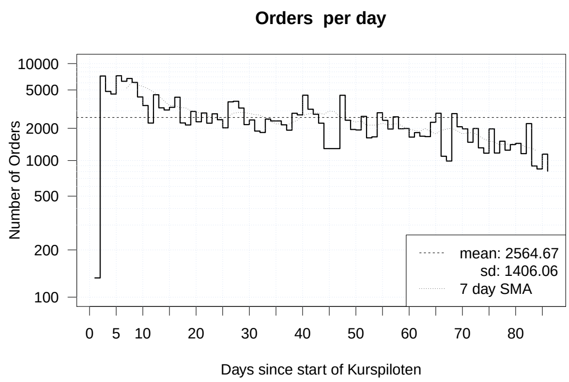
The dataset used is taken form the Kurspiloten market running from 2011-09-02 until 2011-11-25. In total 3,463 participants registered for the Kurspiloten market. 3,217 participants activated their accounts by confirming their email address. Of those, 2,283 submitted at least one order. 1,912 participants submitted at least one order that led to a transaction. During the sample period 144 stocks were paid out. Overall, participants submitted 215,432 orders. Since not every order can be executed against a matching counter-order (e. g., too high/low price), the submitted orders resulted in 131,561 executed transactions. Just 327 of them are submitted via KAPP (cf. Chapter 4.2.2). For every single order the trading value can be calculated and thus, if an order was profitable or not.

Most participants registered in the first few days and thus were able to participate for the majority of the market's runtime (Figure 4.3a). After the first week the registration count per day stabilized around 25 before it dropped to around five in the last two weeks.

Nevertheless participants registered until the last day of the market. The number of orders peaked in the market's first week and stayed above 2,000 orders per day for about two-thirds of the runtime (Figure 4.3b). In the last third of the market lifecycle, it slowly declines towards the minimum point of around 800 orders per day. With more than 2,500 orders submitted on an average per day (Figure 4.3b) the dataset contains 131,561 transactions.



(A) *Registrations per Day*



(B) *Orders per Day*

FIGURE 4.3: *Activity per Day including Simple Moving Average (SMA)*

4.3 Economic Indicator eXchange

The *Economic Indicator eXchange* (EIX)[6] is a prediction market for macro-economic indicators. Initially, the market was designed to forecast economic indices only, as the name indicates. It launched in October 2009 to predict indicators for, e. g., unemployment figures, gross domestic product, and inflation rates (cf. Figure 4.4). Starting in November 2013, the EIX was extended for political indicators (cf. Figure 4.2). This extension was marketed as the *Political Indicator eXchange* (PIX).

4.3.1 Market Design

The work presented in Chapters 6 & 9 is conducted on a prediction market called *Economic Indicator eXchange* (EIX). The goal of EIX is to *"scientificly test and respectively demonstrate*

[6] It was marketed as *Handelsblatt Prognosebörse EIX*.

the possibility of conducting economic forecasts with a prediction market in a field study."[7] The EIX play money prediction market was started in 2009.[8] It is designed to forecast macro-economic variables. In order to reach out to a broad audience interested in economic trends the market was also operated in cooperation with the leading German economic newspaper *Handelsblatt*.

The EIX was planned as a one year project, but was continued in yearly rounds due to its success.[9] Teschner et al. (2011) details the reasoning behind market design for the first two versions of the EIX. Until May 2011 the EIX was run as a web-based system only. In June 2011 a mobile trading application was released called EIX-Market-App (EMA) which provides mobile access to the underlying market system. The goal of the market is to forecast economic indicators up to nine months in advance by continuously aggregating economic information. The market is designed as a continuous double auction without a market maker. After registration participants are endowed with 1,000 stocks of each contract and EIX€ 100,000 (short for 'Economic Indicator eXchange € '). The continuous economic outcomes are represented by one stock and paid out at data release according to a linear payout function as depicted in Table 4.4. To increase participants' motivation and to provide incentives to truly reveal information, in version 3 of EIX prizes worth € 36,000 are offered; eight yearly prizes (total value € 10,000) are awarded according to the portfolio ranking at the end of the market period. In version 4, the value and number of the prices are slightly smaller. Specifically, prizes worth € 4,550 are raffled amongst participants; one yearly prize worth € 299.99, and three prizes per quarter with a total worth of around € 800 in markets dealing with economic indicators as well as more than € 1,000 in political markets (see following paragraph).

Political Markets

In 2013, the EIX was extended to include political markets, called the *Political Indicator eXchange* (PIX). Initially, the PIX was used to collect predictions on the Lower-Saxony state elections. Afterwards, a similar setup attempted to predicted the German federal election

[7] Author's translation; original: *"Ziel des Forschungsprojektes ist es im Rahmen einer Feldstudie die Möglichkeit der Konjunkturprognose mit einer Prognosebörse wissenschaftlich zu testen beziehungsweise zu demonstrieren."* (Jäger-Ambrozewicz, 2009)

[8] The EIX was set up as a joint project. Besides the Institute for Information Systems and Marketing (IISM) at the Karlsruhe Institute of Technology, the EIX project was supported by Forschungszentrum Informatik FZI, Institut der deutschen Wirtschaft Köln, and Verlagsgruppe Handelsblatt.

[9] The work at hand focuses on the years three (2011-11-01 – 2012-10-31) and four (2012-11-05 – 2013-10-29) of the EIX. The years one (2009-11-30 – 2010-10-31) and two (2010-10-01 – 2011-10-31) are covered as far as necessary. For an in-depth analysis of the first two year of EIX see Teschner (2012).

in late 2013. Specifically, two markets are run in order to gather predictions: The *candidate* market is comprised of *Winner-takes-all* contracts (cf. Section 3.2) representing the chances of being elected for Chancellor candidates respectively Minister-President candidates. The *party* market contains different *Index* contracts (cf. Section 3.2), one for each promising party as well as a rest-of-field contract. These markets run on the EIX market system but are separate from the economic indicators. Trading takes place in a separate play-money currency called 'PIX€'. Additionally, a separate ranking for political markets is provided.

The German federal voting system, in which each voter has two votes, is rather complex. It makes use of *proportional representation* (PR) and *method of majority decision* (MD) and works roughly like this: the *Erststimme* (first vote) determines in each constituency which delegate is sent to the *Bundestag* (parliament) using MD; the *Zweitstimme* (second vote) uses PR to determine the share of seats each party achieves. If no party is able to gain the bare majority of votes, elected parties start exploratory talks in order to form a coalition. Afterwards, the *Bundestag* (i. e., all delegates of the parliament) elects the Chancellor on a proposal of the *Bundespräsident*. The Lower-Saxony state election system is quite similar to the aforementioned voting system. Put simply, voters elect representatives sent to the *Landtag* (parliament) using PR, which subsequently elects the Minister-President using MD.

The Lower-Saxony state election market operated from 2012-11-05 to 2013-01-20.[10] The *candidate* market contains four *Winner-takes-all* contracts (cf. Section 3.2) representing David McAllister, Stephan Weil, and rest-of-field. In the *party* market, seven *Index* contracts are used to represent the parties CDU, SPD, FDP, Grüne, DIE LINKE, Piraten, and rest-of-field.[11] Afterwards, an analogous market design is used for the German federal elections. It operated from 2013-01-23 until just before the election on 2013-09-22.[12] A quite long runtime of eight months allows plenty of time for collecting data. The German federal elections market predicts the election outcome continuously from 2013-01-23 until the election on 2013-09-22.

For the German federal elections, the *candidate* market is comprised of four *Winner-takes-all* contracts (cf. Section 3.2) belonging to Chancellor candidates Dr. Angela Merkel, Peer Steinbrück, an unnamed Green candidate (placeholder) and rest-of-field. The *party* market contains eight different *Index* contracts, for CDU/CSU, SPD, FDP, Grüne, DIE LINKE,

[10]The market closed at 07:45 o'clock on 2013-01-20, as the election started the same day at 08:00 o'clock.
[11]The entire names of the abbreviated parties can be found in the List of Abbreviations.
[12]The market closed at 07:45 o'clock on 2013-09-22, as the election started the same day at 08:00 o'clock.

Piraten, AfD[13], and rest-of-field.

The best performing traders by portfolio value win prizes after the market closes. To account for percentage values, *Winner-takes-all* contracts pay PIX € 100 if the respective candidate becomes the next Minister-President of Lower-Saxony respectively the next German Chancellor, and *Index* contracts pay their respective parties' election result percentage in PIX € . Registration is free, but every person is only allowed one trading account. Upon registration, traders receive an initial endowment of PIX € 100,000 and 1,000 stocks of each contract. Since the contracts in each market are interdependent due to their different underlying events, this constraint dictates that prices should sum up to PIX € 100 (which corresponds to 100 %). To easily enforce this, traders have the possibility to buy and sell the unit portfolio, consisting of one of each contract in a market, for PIX € 100. Therefore, if the sum of best bids equals up to over PIX € 100, or the sum of all best asks equals up to less than PIX € 100, there is opportunity for arbitrage. For an in-depth discussion of bundle trading see "Basic Portfolios" in Luckner and Weinhardt (2008, pp. 55). The remaining market details are analogous to the EIX market (cf. Section 4.3): Participants can submit limit orders continuously – with the exception that short sales are not allowed. As limit orders with an extreme price can be used, there is no need for market orders. Orders are matched continuously according to the order precedence rule. The five best bids and asks for each contract are displayed in an order book.

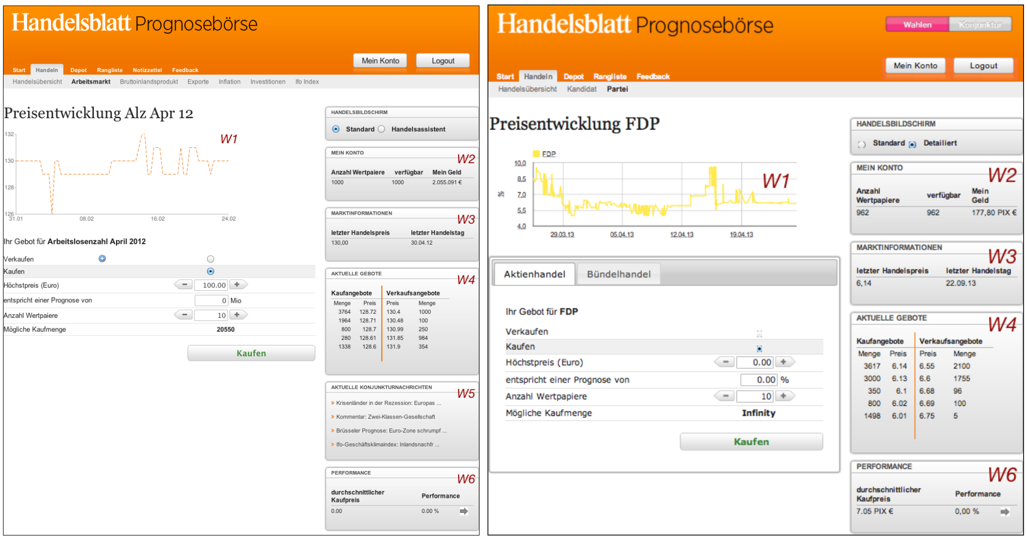
4.3.2 Market Interfaces

EIX Web Interface

The web-based trading interface of version 3 of the EIX is displayed in Figure 4.4a.[14] Figure 4.4b depicts the web-based trading interface of the PIX, which was introduced in version 4 of the EIX. In both interfaces, participants have convenient access to the price development (W1), the account information (W2), market information (W3) such as the last trading day, the order book with five levels of visible depth (W4). As additional information, Handelsblatt provides access to an up-to-date economic news-stream (W5) which

[13]In April 2013, the rest-of-field contract rose over 10 %. An integrated survey (as presented in Chapter 9) was used to gather insights about that development from the market participants. As a majority suspected the newly founded party AfD to be behind that increase, it was decided to split the rest-of-field contract in a new rest-of-field contract and an AfD contract in 2013-04-30.

[14]As the trading interface of EIX did not change notably between version 3 and version 4, no dedicated screenshot of version 4 is contained.

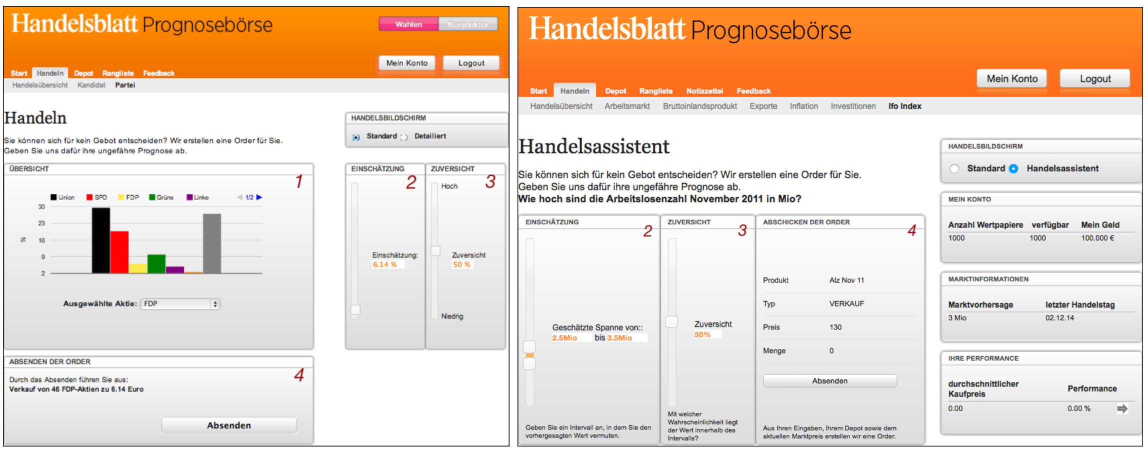


(A) *EIX Web Trading Screen (Version 3)* (B) *PIX Web Trading Screen (Version 4)*

FIGURE 4.4: *EIX and PIX Web Trading Screens (annotated)*

was included in the EIX interface only. Furthermore, the trader's average purchase price of the current indicator (W6) is displayed.

The trading wizard for the PIX is depicted in Figure 4.5a. First, all tradable contracts are shown in an overview chart (1), where their current prediction in the market is shown as a bar plot. By clicking on the specific bar, traders can choose the indicator to trade.



(A) *PIX Web Trading Wizard (Version 4)* (B) *EIX Web Trading Wizard (Version 3)*

FIGURE 4.5: *EIX and PIX Web Trading Wizard Screens (annotated)*

Additionally, the indicator can be chosen from a drop-down list below the overview chart. Second, traders report their estimation of the election outcome with the help of a slider

(2). The resulting percentage is used to derive a limit price for the order suggested. Third, traders report their confidence via another slider (3). This value is used to determine the quantity of stocks configured in the resulting order. Finally, the suggested order is displayed (4) and traders can submit these with the clicking of a button. Figure 4.5b depicts the trading wizard for the EIX in version 3.[15] It differs slightly from the PIX trading wizard in two aspects. First, it does not offer the possibility to chose the product to trade inside the trading wizard itself. This is a consequence of the defined work flow for trading economic indicators, which lets users chose the product to trade in a separate screen. Second, instead of picking an exact expectation, traders can set their expectation's upper and lower bounds (2). Subsequently, traders set their confidence level (3), before they are able to submit the derived order (4).

EIX Mobile Interface

The EIX-Market-App (EMA) is a mobile client for the EIX designed for Apple's iPhone. EMA offers all of EIX's core features, i. e., submit and cancel orders, checking one's holdings in the portfolio, additional information like the order book, news, and so forth. EMA's frontend is a compromise of two design goals. First, EMA was intended to be easy to use for new users. Second, existing EIX-users should be able to use the App with minimal learning effort. Due to the limited screen size of the iPhone, it is not reasonable to let EMA look exactly like EIX's web-interface. EMA's frontend tries to be close to EIX's web-interface by sharing the same menu-structure and nomenclature (Figure 4.6).

Analogous to EIX's web-interface, EMA offers six stock related information screens linked from the trade screen (Figure 4.6b). To allow research about users' information usage prior to submitting an order, the consumption of the six information panels (respectively screens, in case of EMA) are logged in both IS. Both EIX and EMA track the time a user needs to create and submit an order as well as the information used in this process.

[15] As the trading interface of EIX did not change notably between version 3 and version 4, no dedicated screenshot of version 4 is contained.

(A) *upper part* (B) *lower part*

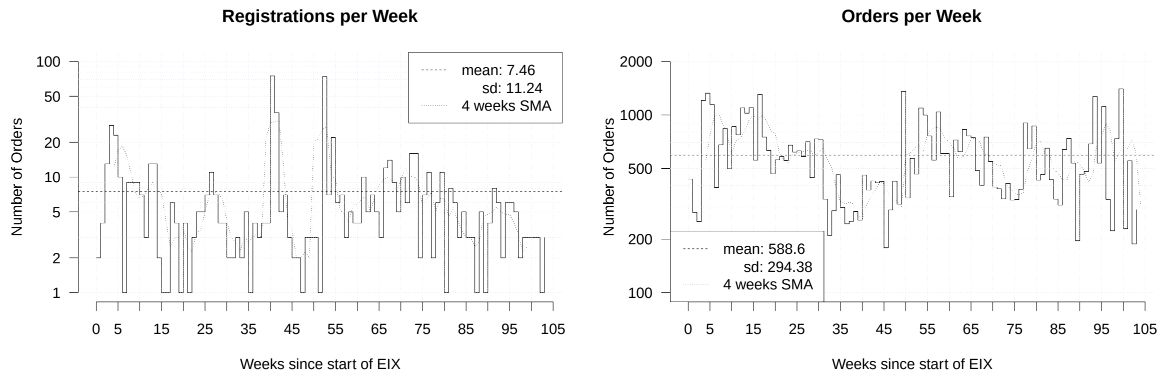
FIGURE 4.6: *EIX Mobile Trading Screen*

4.3.3 Dataset and Descriptive Statistics

EIX

Figure 4.7 depicts logarithmized activity measures for versions 3 & 4 of the EIX market from 2011-11-01 until 2013-10-29.[16] Specifically, Figure 4.7a shows the number of persons registering for EIX aggregated on a weekly basis. Furthermore, this Figure contains a monthly simple moving average of the registration figures. On average more than seven persons registered per week. Figure 4.7b depicts the number of orders submitted per week besides a monthly simple moving average. Nearly 600 orders were submitted per week on average.

[16]These charts do not contain data related to the political stock market PIX. Instead, only orders concerning economic indicators as well as registered persons that traded those products are considered.

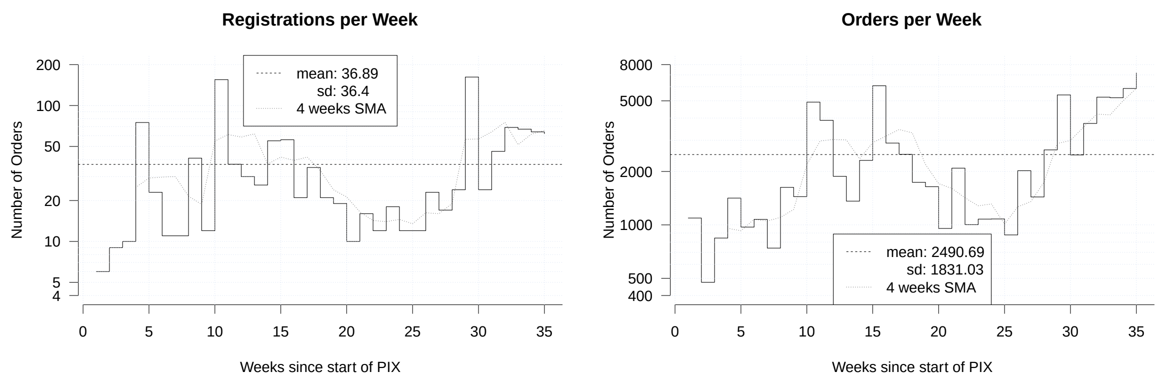


(A) *Registrations per Week* (B) *Orders per Week*

FIGURE 4.7: *EIX Activity per Week including Simple Moving Average (SMA)*

PIX

Figure 4.8 displays the logarithmized time series of the number of registered persons and submitted orders for the PIX market. Both charts show weekly aggregated data in addition to a four weeks simple moving average. As shown in Figure 4.8a, on average more than 36 persons registered per week. Here, participants submitted on average about 2,500 orders per week as depicted in Figure 4.8b.



(A) *Registrations per Week* (B) *Orders per Week*

FIGURE 4.8: *PIX Activity per Week including Simple Moving Average (SMA)*

4.4 Continuous Market Engineering – Timeline

This section gives an overview of the changes carried out on the beforehand introduces prediction markets.

TABLE 4.2: *Operational Timeline of Regarded Prediction Markets*

Market	Area of Prediction	Version	Opening Date	Closing Date
EIX	Economic Indicators	1	2009-10-30	2010-10-31
EIX	Economic Indicators	2	2010-10-01	2011-10-31
EIX	Economic Indicators	3	2011-11-01	2012-10-31
EIX	Economic Indicators	4	2012-11-05	2013-10-29
PIX	Lower Saxony State Election	4.1	2012-11-05	2013-01-22
PIX	German Federal Election[a]	4.2	2013-01-23	2013-09-22
Kurspiloten	Stock Indices and Commodities	1	2011-09-02	2011-11-25

Notes: [a]Introduction of contract 'AfD' in German federal election market took place on 2013-04-30.

Table 4.2 depicts the runtime of the prediction markets alongside their major versions. The first version of EIX started to operate in late 2009. Teschner (2012) describes the first and second version of EIX in detail. The work at hand focuses on version 3 & 4 of EIX, its submarket PIX, and Kurspiloten. These operated between September 2011 and November 2013.

TABLE 4.3: *Timeline of Conducted Studies*

Study	Chapter	Market	Begin	End
Trader's Market Predisposition[a]	5	Kurspiloten	2011-09-02	2011-12-14
Reading a Trader's Mind[b]	6	EIX (v4.2)	2013-06-21	2013-09-23
Stationary vs. Mobile	7	Kurspiloten	2011-09-02	2011-11-25
Interface/Disposition Effect	8	Kurspiloten	2011-09-02	2011-11-25
Survey Comparison	9	EIX (v4.2)	2013-07-29	2013-08-25

Notes: [a]Market data from 2011-09-02 until 2011-11-25 was considered. The follow-up questionnaire was conducted between 2011-11-30 and 2011-12-14; a reminder was sent out on 2011-12-06; [b]A translated version of the question *"Which party can you identify with the most, when it comes to national politics?"* (cf. Sjöberg, 2009) was asked at two different periods (2013-06-21 until 2013-07-31 and 2013-08-26 until 2013-09-23).

Table 4.3 lists the studies conducted in the subsequent chapters along with the associated prediction market instance. Table 4.3 relates the studies conducted in the work at hand to the specific market they are based upon. Furthermore, the timeframe in which data was collected can be found in the aforementioned table. Additional details can be found in the Table notes.

TABLE 4.4: *Timeline of Tradable Contracts on EIX Market*

				Version			
Indicator	Unit	Cycle	Payout Function	1	2	3	4
Export	%	monthly	$100+\alpha\times(\frac{I_t-I_{t-1}}{I_{t-1}})$	✓			
Export	Bil. €	monthly	$30+\frac{ABS(I_t)}{10^9}$		✓	✓	✓
Gasoline	€	monthly	$ABS(I_t)$				✓
GDP	%	quarterly	$100+\alpha\times(\frac{I_t-I_{t-3}}{I_{t-3}})$	✓	✓	✓	✓
Inflation	%	monthly	$100+\alpha\times(\frac{I_t-I_{t-12}}{I_{t-12}})$	✓	✓	✓	✓
Ifo Index	%	monthly	$100+\alpha\times(\frac{I_t-I_{t-1}}{I_{t-1}})$	✓			
Ifo Index	Points	monthly	$ABS(I_t)$		✓	✓	✓
Investments	%	quarterly	$100+\alpha\times(\frac{I_t-I_{t-1}}{I_{t-1}})$	✓	✓	✓	
Unemployment	Num.	monthly	$100+\frac{ABS(I_t)}{10^6}$	✓	✓	✓	✓

Notes: I: Indicator; t: time in months; α=10

Tables 4.4 & 4.5 are compilations of all products that where tradable on the EIX and PIX during their runtime. Table 4.4 lists all economic indicators that were tradable throughout the first four versions of EIX.

TABLE 4.5: *Tradable Contracts on PIX Market*

Indicator	Unit	Contract Type	Payout Function
Candidates LSSE[a]	%	Winner-takes-all	PIX€ 100 if C is elected, otherwise PIX€ 0
Party LSSE[a]	%	Index	ABS(Vote Share$_P$)
Candidates GFE[b]	%	Winner-takes-all	PIX€ 100 if C is elected, otherwise PIX€ 0
Party GFE[b,c]	%	Index	ABS(Vote Share$_P$)

Notes: C: Candidate; P: Party, t: time in months; [a]LSSE: Lower-Saxony State Election 2013; [b]GFE: German Federal Election 2013; [c]Introduction of contract 'AfD' in German federal election market on 2013-04-30

Basically, there are two points in time at which indicators, respectively their payout function, were changed. First, between version 1 and version 2, the payout function of the indicators 'Export' and 'Ifo Index' were modified from a relative measure with scaling towards an absolute representation.[17] In case of 'Export', the payout function was additionally scaled and added to a fix amount. Second, starting with version 4, the indicator 'Investments' was canceled due to low activity in favor of 'Gasoline', which was expected to be (i) easy to understand and (ii) more strongly connected to participants' daily routine than its predecessor.

Finally, Table 4.5 list the four different indicator classes used in PIX besides their payout functions. The timeframe, in which indicators of each classes were tradable resemble the runtime of the PIX as depicted in Table 4.2. The candidate stocks are modeled as Winner-takes-all contracts, which means that they pay PIX€ 100 in case the candidate represented by this stock is elected Chancellor respectively Minister-President; otherwise they are worthless. In contrast, Index contracts are used to model the party stocks. Hence, they pay vote share of their underlying party in PIX€ at the time of release of the election results.

4.5 Summary

This chapter presented the prediction markets used in the work at hand. First, the Kurspiloten market, used in Chapters 5, 7 & 8 is introduced. Second, the Economic Indicator eXchange with its submarket, the Political Indicator eXchange is described. It builds the basis for the studies described in Chapter 6 & 9. Besides their web-based user interfaces both markets also provided mobile user interfaces in forms of specifically designed mobile applications. A major distinction lies in the markets' durations; whilst Kurspiloten has a rather short runtime of 12 weeks, each version of EIX operated for about one year. More specific descriptive data on the prediction markets described can be found in the respective chapters.

[17] See Teschner (2012) for a detailed discussion on this topic.

Part III

Insights from Continuous Market Engineering

Chapter 5

Analyzing *Agent Behavior*: Assessing Trader's Market Predisposition

> " In economics, the majority is always wrong."
>
> JOHN K. GALBRAITH

5.1 Introduction

IN this study behavioral aspects of market participants are linked with the quality of their trading decisions and behavior in the market. Creating a link between behavioral aspects of the participants and quality is important in that the quality of the predictive power is directly negatively affected if participants make systematically biased decisions. This is a relatively well known, but still not well understood or studied hypothesis of behavioral finance literature. In *Kurspiloten* market *decision quality* is obviously described by the participants' trading performance as well as their share of profitable trades. Current research does not clearly answer the question which personal attributes support or hinder specific successful behavior in markets – and maybe never will. The current approach is extended by taking *user heterogeneity aspects* such as personal attributes into account. Besides trading performance it focuses on trader activity and whether they provide or take liquidity to/from the market as qualitative measures for *trading behavior*. Specifically, a two-staged study is conducted to investigate the influence of *cognitive reflection* abilities, grade of *risk aversion*, and use of *emotion regulation strategies* on trading behavior and decision quality in a prediction market context.

The remainder of this chapter is structured as follows: Section 5.2 gives a review of related work on personal attributes, risk aversion, and trading behavior. The experimental setting and research questions are presented in Section 5.3. Section 5.4 describes the dataset and the methodology used. Subsequently, the results are presented from two perspectives: *trading behavior* and *decision quality*. Finally, Section 5.5 concludes this chapter.

5.2 Related Work

5.2.1 Personal Attributes and Trading Behavior

Psychologists have demonstrated a variety of systematic departures from 'rational' decision-making by individuals. These lead to substantial information processing biases or judgment biases and colored expectations (Forsythe et al., 1999). Markets suffer from biases as well and there is an ongoing debate to which extent their efficiency is affected (Arrow et al., 2008). Objectively irrelevant (Huber et al., 2008) and selectively presented information (Dittrich et al., 2005) can and does influence individual trading behavior. A promising approach to describe and explain financial decision-making may be the explicit consideration of psychological factors. Lo et al. (2005) for example have shown the negative influence of extreme emotional states on trading performance. Additionally, they conclude that *"[t]he lack of correlation between personality traits and trading performance begs for additional data and a more refined analysis [...]"* (Lo et al., 2005). Their approach of acquiring psychological factors via personality questionnaires seems promising. Frederick (2005) introduced a well-established questionnaire to measure cognitive ability, the *cognitive reflection test* (CRT). It builds upon the existence of two types of cognitive processes which Stanovich and West (2000) call "System 1" and "System 2" processes. *"System 1 processes occur spontaneously and do not require or consume much attention. [...] System 2 processes [are] mental operations requiring effort, motivation, concentration, and the execution of learned rules"* (Frederick, 2005). By offering participants three short tasks, which – at first glance – seems to be solved best by "System 1" processes while actually being more complex tasks (i. e., "System 2"), it is possible to differentiate the more impulsive from the more cognitive reflective ones. The *ten paired lottery* (TPL) introduced by Holt and Laury (2002) is a widely used risk aversion test that offers *"[a] menu of paired lottery choices[,] structured so that the crossover point to the high-risk lottery can be used to infer the degree of risk aversion"* (Holt and Laury, 2002). Participants can choose between 'A' (safe choices) and the more risky 'B'. By design, the risk neutral choice pattern is four 'A' choices followed by six 'B' choices. Gross and John (2003) introduced a questionnaire to determine emotion regulation strategies, the *emotion regulation questionnaire*. It consists of ten statements – four concerning suppression and six concerning reappraisal – the participant agrees or disagrees with on a seven-point Likert scale. The concept of *reappraisal* takes place in the context of antecedent-focused emotion regulation and means a cognitive change in the interpretation of a situation. *Suppression* happens in the context of response-focused emotion regulation and aims to hide a specific emotion. All three questionnaires are rather short whilst reliable and can therefore be used altogether in one questionnaire without overly stretching a participant's attention.

5.2.2 Risk Aversion and Trading Behavior

Several authors have identified risk aversion as a reason for certain market behavior (e. g., Subrahmanyam, 1991). It may cause participants to not make profitable but risky trades in a market. If participants suffer from this aversion, valuable information may not be impounded into prices and thereby reduce the predictive power of a market. Unfortunately, useful insights can only rarely be obtained from empirical data on security prices since risk aversion measures must be obtained independently of trading data. By merging household investment decisions with data from external risk questionnaires Wärneryd did not find a relationship between risk aversion and portfolio choice (Wärneryd, 1996). This is in line with findings from an empirical asset market in which participants' portfolio choice is unrelated to a risk aversion proxy (Güth et al., 1997). In contrast to portfolio choice, individual market behavior seems to be influenced by risk aversion. Fellner and Maciejovsky (2007) find that the higher the degree of risk aversion, the lower the observed market activity. Kirchler and Maciejovsky (2002) find the higher the degree of risk aversion the lower the total number of contracts traded. In an early experimental study, Ang and Schwarz (1984) separated participants in two markets according to their degree of risk aversion. They show that the market with lower risk aversion (speculators) exhibit greater volatility but it also tend to converge closer and faster to the expected equilibrium price than the risk averse (conservative) market. Finally, the interaction between risk attitude and overconfidence with respect to trading activity deserves further attention. Theoretical finance models predict higher market activity as a consequence of overconfidence[1] (Barber and Odean, 2001). People tend to be overconfident about their capabilities and level of knowledge. This could also negatively impact the information content of prices.

5.2.3 Trading Behavior in the Market

Trading behavior in the market is measured via two measures. First, the traders' activity is used; i. e., the number of submitted orders. Second, their 'roles' in the market are regarded. A common perspective to categorize *trading behavior* is to group traders depending on how they submit their orders. One possibility is to separate between a) *liquidity providers* or *market makers* and b) *liquidity takers* or *price takers*. *Market makers* usually buy and sell the same contract at the same time, trying to profit from the spread. Another feature is placing orders on top of the order queue instead of taking the opposite first offer. The *marginal trader hypothesis* by Forsythe et al. (1992) assumes that marginal and not average traders determine prices. These traders 'make the market' and appear to be more rational (Oliven and Rietz, 2004), plus they are more unlikely to produce trading violations (Forsythe et al., 1992). Oliven and Rietz (2004) report that *price takers* make errors on average 47 % of the time whereas *market makers* had an average 8 % error rate. Consistently Forsythe et al. (1999) describe an error rate for *price takers* as high as nearly 6 times the error rate for *market makers*. As a result when traders act as *market makers*, they make fewer mistakes and hence appear more rational. Furthermore *market makers* serve as liquidity providers and allow continuous trading (Luckner, 2008). The usually small group of *market makers* has a disproportionately large effect on aggregated market behavior (Sunstein, 2006). Previous work on *trading behavior* consistently suggests that *liquidity providers* perform better in market environments. In order to understand the motives behind the self-selection into these roles, Oliven and Rietz (2004) use demographic information. They find that "*[. . .] this choice is significantly affected by market-specific experience and general financial knowledge, education, sex, and religious affiliation [. . .]*" but nevertheless "*[. . .] remains largely unexplained*" (Oliven and Rietz, 2004).

5.2.4 Service Analytics

Following Fromm et al. (2012, p.143), service analytics can be classified in two dimensions. First, complexity can be separated in *basic analytics* as a foundation (comprising

[1] Overconfidence refers to the habit of overestimating one's ability to perform a task.

data management and reporting) and *advanced analytics* using methods from statistics and operations research building on top of it. Especially the latter is predestined to unveil a service's full potential. Second, the analytics' scope can be ranged in provider, encounter, or customer data. In an e-service system context such as a prediction market, the customer data to apply *advanced service analytics* can often easily be obtained, since the provider and customers are connected by design (Fromm et al., 2012).

5.3 Setting and Research Questions

5.3.1 Experimental Setting

In order to answer the research questions presented in the next Subsection (5.3.2) a two-staged field study on the Kurspiloten market (cf. Section 4.2) is conducted. In a first phase, participants took part in an online prediction market. For the study's second phase all market participants are invited to take part in a four-section online survey, five days after the market's end. The first section concerns general feedback of the market platform and its game design. The main part combines the three questionnaires introduced in Subsection 5.2.1. Namely, the *cognitive reflection test*, followed by the *emotion regulation questionnaire*, and last the *ten-paired lottery*. The questionnaire closes with a "final evaluation question" which asks if participants answered truthfully throughout the questionnaire. The survey was active for 14 days and participants were incentivized by giving away ten Amazon vouchers worth € 30 each via a raffle.

5.3.2 Research Questions

Advanced service analytics is applied on an e-service system in order to gain comprehensive insights on customers' *market predisposition*. Based on this, it should be possible to substantially improve a customer's service experience in a second step. This can be achieved by adapting the service to customers' preferences and abilities via personalized tweaks such as interface adaptions, and product choice. In particular, an attempt is made to shed some light on the research questions 1 (following an aspect of Oliven and Rietz (2004)) and 2 presented in Subsection 1.2, which are:

> **Research Question 1:** *How do selected personal attributes (RA, CRA, and ERS) influence trading behavior in markets?*

Research Question 2: *How do selected personal attributes (RA, CRA, and ERS) influence decision quality in markets?*

Specifically, this analysis focuses on the influence of *cognitive refection* abilities, *risk aversion*, and *emotion regulation strategies* on the aforementioned *trading behavior* and *decision quality*. As a person's cognitive reflection is known to be positively correlated to her IQ as well as other measures for cognitive ability (Frederick, 2005), it is assumed that a higher cognitive reflection leads to 'better decisions' in general. For trading behavior, – in particular for activity – it is not quite clear, what 'better' means, nevertheless it is expected that the more cognitive reflective traders are less likely liquidity takers (i. e., price takers) (Forsythe et al., 1999). In case of decision quality, it is expected that a high CRT-value leads to a higher trading performance as well as a higher probability to make a profit. As stated in Section 5.2, risk aversion has been shown to have an impact on trading behavior. Hence, it is expected that risk averse traders are less active. Furthermore, it is assumed that risk attitude does induce certain trading behavior. According to a study of Fenton-O'Creevy et al. (2011), the emotion regulation strategy used by traders differs according to their experience and performance. Therefore, it is expected that certain behavioral patterns depending on the emotion regulation strategy used are discovered. Among others, these behavioral patterns include how traders engage in a market, what their decision quality will be or how they self-select into the roles of *price takers* or *market makers*.

5.4 Results

In this section the empirical findings are presented, starting with descriptive statistics. Subsequently, trading behavior in the market and the traders' decision quality is analyzed.

5.4.1 Descriptive Statistics and Methodology

In total, 512 at least partly processed online questionnaires are received; 386 of them are completely filled. 320 of those contain a positive answer to the "final evaluation question". The median processing time of those 320 replies is 11 minutes 26 seconds (mean: 26m 43s) for the whole questionnaire and 9 minutes 21 seconds (mean: 24m 26s) for the main part containing CRT, ERQ, TPL. In order to statistically analyze this dataset the survey responses have to be filtered and – as well as the trading data – operationalized. Therefore, replies are filtered based on the answer for the TPL. The so-called "stay in bed" types (i. e., participants that report to be irrational risk averse) are filtered. These respondents have chosen 'A' over 'B' in question nine and/or ten of the TPL, where the expected payoff is lower for 'A' than for 'B' ($ 1.96 vs. $ 3.47 and $ 2 vs. $ 3.85). Note, that the so-called "ABBA" types of the TPL (i. e., respondents who switch multiple times forth and back between A and B) are not filtered. According to Holt and Laury (2002) *"[e]ven for those who switched back and forth, there is typically a clear division point between clusters of A and B choices, with few 'errors' on each side. Therefore, the total number of 'safe' A choices will be used as an indicator of risk aversion."* In this survey, the mean difference in the number of 'A' answers with and without "ABBA" types is a mere 0.03 (4.86 to 4.89). After this step, 246 questionnaires are left. Of those, 50 participants did not actively trade in the Kurspiloten market, i. e., they submitted no order at all. This leaves 196 usable questionnaires for evaluation. This corresponds to 10.25 % of active participants (50.78 % of completely filled questionnaires) or to an overall response rate of 20.19 % (completely filled questionnaires in relation to active participants), which is a fairly normal response rate for online questionnaires (e. g., Cook et al., 2000; Ranchhod and Zhou, 2001; Deutskens et al., 2004). The variables used in the analyses (Table 5.1) are described in the following.

TABLE 5.1: *Variables*

Variable	Description	Value
CRT_{high}	Cognitive Reflection Test – Three correct answers = 1	1 or 0
$TPL_{risk\ averse}$	Ten Paired Lottery – Five or more 'safe choices' = 1	1 or 0
$ERQ_{suppress}$	Emotion Regulation Questionnaire – Suppression is used = 1	1 or 0
$ERQ_{reappraise}$	Emotion Regulation Questionnaire – Reappraisal is used = 1	1 or 0
buy_o	Order o is a buy = 1	1 or 0
$initialize_o$	Order o initializes a trade = 1	1 or 0
$quantity_o$	Size of order o in stocks	[1, inf]
limit $price_o$	Limit price of order o	[.01, inf]
$profit_o$	Profit made with order o	[-inf, inf]
win_o	Order o was profitable = 1	1 or 0
order $count_p$	Number of orders executed by participant p	[1, inf]

The CRT consists of three questions that can be answered either correctly or incorrectly. To derive a dichotomous variable for the CRT, participants with zero to two correct answers are assigned into the group $\mathrm{CRT}_{\mathrm{low}}$; thereby only participants who answered all three questions correct are put in the group $\mathrm{CRT}_{\mathrm{high}}$. Since the responses of the ERQ are collected via a seven point Likert scale, the mean of the answers concerning the suppression and reappraisal strategy are calculated separately and normalized to the interval $[-1, 1]$. Finally, 1 is assigned to the dummies $\mathrm{ERQ}_{\mathrm{suppress}}$ or $\mathrm{ERQ}_{\mathrm{reappraise}}$ if the normalized averages of replies concerning the corresponding strategy are greater or equal zero; else they are set to 0. The reliability of the ERQ is estimated with Cronbach's α (Cronbach, 1951), which is 0.673 for the $\mathrm{ERQ}_{\mathrm{suppress}}$ questions and 0.819 for the $\mathrm{ERQ}_{\mathrm{reappraise}}$ questions. With an α of more than 0.8, the assessment of 'Reappraisal' can be considered good. Although, the α for $\mathrm{ERQ}_{\mathrm{suppress}}$ is slightly below 0.7, the survey's results can be considered reliable since the latent construct 'Suppression' is measured – by design of the ERQ – with just four items. Since, by design, Cronbach's α rises with N, the aforementioned α value is in an acceptable range for a four-item construct. Responses of the TPL are also segregated into two groups: $\mathrm{TPL}_{\mathrm{risk\ averse}}$ is set to 1 for participants with five or more 'A' choices, while it is 0 for participants with four or less safe choices. The trading direction is identified by the variable *buy*, which is 1 for a buy and 0 for a sell order. The variable *initialize* is used to distinguish between liquidity taking and liquidity providing orders. An order that is not immediately executed provides liquidity to the market, whereas an order that initializes a trade directly after submission to the market 'takes' liquidity from it. (For example, a buy order a of 125 Stocks for P€ 120.00 is submitted while a sell order b of 100 stocks for P€ 120.00 and another sell order c of 150 Stocks for P€ 119.95 are the highest sell orders in the order book. The initializing order is order a, since it initializes the trade, as it completely fulfills order b and partly (25 units) order c. Note, that under certain circumstances order b and c can also be initializing, due to a prior (partly) execution.) In the first case *initialize* is set to 0, since the order does not trigger a trade, else it is set to 1, i. e., if an order takes liquidity from the market. Furthermore, the limit price of an order in P€ (*limit price*) and the number of shares traded (*quantity*) is used. The variable win indicates if a specific transaction led to a (positive) *profit*. Last, the number of orders submitted per participant is encoded by *order count*.

Two types of regression models are used to analyze the dataset. First, OLS regressions are used to estimate the personal attributes' influence on participants' trading activity (Equation 5.1) as well as on the profit per trade (Equation 5.3). Second, logistic regressions are used to investigate participants' trading strategy (Equation 5.2) and profitability

(Equation 5.4).

$$\text{order count}_p = i \quad +\beta_1 \times CRT_{high} + \beta_2 \times TPL_{riskaverse} \\ +\beta_3 \times ERQ_{suppress} + \beta_4 \times ERQ_{reappraise} \tag{5.1}$$

Equation 5.1 connects participants' personal attributes to participant's market activity proxied by the number of submitted orders. The personal attributes regarded comprise *cognitive reflection abilities*, *risk aversion*, and *emotion regulation strategies*.

$$log(\frac{\pi_{initialize}}{\pi_{Trade}}) = i \quad +\beta_1 \times CRT_{high} + \beta_2 \times TPL_{riskaverse} \\ +\beta_3 \times ERQ_{suppress} + \beta_4 \times ERQ_{reappraise} \\ +\beta_5 \times \sum_{i=1}^{12}(\gamma_i \times M_i) \\ +\beta_6 \times buy_o + \beta_7 \times quantity_o + \beta_8 \times limit\_price_o \tag{5.2}$$

Equation 5.2 examines the trading strategy used (*liquidity providing* vs. *liquidity taking*) whilst considering personal attributes on a per-order level. Furthermore, it is controlled for different markets ($M_i, i = \{1, \ldots, 12\}$, cf. Table 4.1), trading direction (*buy*), order size (*quantity*), and price (*limit_price*).

$$profit_o = i \quad +\beta_1 \times CRT_{high} + \beta_2 \times TPL_{riskaverse} \\ +\beta_3 \times ERQ_{suppress} + \beta_4 \times ERQ_{reappraise} \\ +\beta_5 \times \sum_{i=1}^{12}(\gamma_i \times M_i) \\ +\beta_6 \times buy_o + \beta_7 \times initialize_o \tag{5.3}$$

Equation 5.3 relates the profit gained per order in P€ to personal attributes on a per-order level. Again, this model controls for product-specific effects (M_i) and trading direction. Additionally, it is controlled for liquidity providing/taking trading behavior (*initialize*).

$$log(\frac{\pi_{win}}{\pi_{Trade}}) = i \quad +\beta_1 \times CRT_{high} + \beta_2 \times TPL_{riskaverse} \\ +\beta_3 \times ERQ_{suppress} + \beta_4 \times ERQ_{reappraise} \\ +\beta_5 \times \sum_{i=1}^{12}(\gamma_i \times M_i) \\ +\beta_6 \times buy_o + \beta_7 \times quantity_o + \beta_8 \times limit\_price_o + \beta_9 \times initialize_o \tag{5.4}$$

Equation 5.4 estimates the probability to submit a profitable order (in relation the final outcome) as a function of personal attributes. Again, it is controlled for product-specific effects (M_i) and trading direction. Furthermore, order size (*quantity*), price (*limit_price*), and liquidity providing/taking trading behavior (*initialize*).

The results of this study are presented in the following subsections. First, the trading behavior in the market is analyzed, then the traders' decision quality is regarded.

5.4.2 Trading Behavior

In the following the trading behavior in the market is investigated with two types of regression models. First, the traders' personal attributes (namely, *cognitive reflection* abilities, *risk aversion*, and *emotion regulation strategies*) are connected with their *activity* proxied by the *order count* on a per-user basis. Second, the personal attributes are investigated in terms of *trading strategy* on a per-order basis. Therefore, orders are classified as *liquidity providing* or *liquidity taking* via the variable *initialize*.

Activity

A linear regression model built to analyze how the activity per user depends on the personal attributes. As both models in Table 5.2 show, the activity is significantly higher for participants with a high cognitive reflection (standardized coefficients: .16 in A1, .15 in A2). Also, risk aversion significantly increases the number of orders submitted to the market (Model A1, std. coef.: .15). Even though risk aversion is significant positive correlated with the number of submitted orders (cor = .148, t-stat = 2.08), risk aversions' influence declines and its significance fades to the 10 %-level when the participants' emotion regulation strategies are included (Model A2, std. coef.: .13). Similarly, $\text{ERQ}_{\text{suppress}}$ is significantly negatively correlated with the number of submitted orders (cor = −.146, t-stat = −2.07); both emotion regulation strategies have no significant effect in model A2 (using the logarithmized order count as dependent variable leads to similar results.). In contrast to Fellner and Maciejovsky (2007) and Kirchler and Maciejovsky (2002), a robust relation between risk aversion and activity have *not* been found.

Trading Strategy

In order to investigate the participants' personal attributes on their *trading strategy* (i. e., their role in the market), a logistic regression on the variable *initialize* is conducted. The

TABLE 5.2: *Activity*

Model	**A1** order count	**A2** order count
CRT_{high}	165.95**	160.50**
	(2.25)	(2.15)
$TPL_{risk\ averse}$	157.76*	138.89·
	(2.17)	(1.89)
$ERQ_{suppress}$		−118.98
		(−1.62)
$ERQ_{reappraise}$		45.03
		(.57)
(Intercept)	21.97	64.41
	(.35)	(.68)
Adj. R^2	3.69 %	4.05 %
Num. obs.	196	196

Notes: t-statistics in parenthesis; $^{\cdot}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

results in Table 5.3 show that high cognitive reflection favors the initialization of trades (Marginal Effects (mfx): .02 in I1 and I2, .01 in I3) whereas risk aversion hinders it (mfx: −.02 in I1, −.03 in I2 and I3). While the strengths of those effects are about equal in I1, they diverge more and more from I2 to I3. Like risk averse traders, ones using the suppression strategy also tend to initialize less often (mfx: −.02 in I2, −.03 in I3) whereas traders using the reappraisal strategy tend to initialize trades more often (mfx: .04 for I2, .04 for I3). Although the *limit price* of an order is highly significant, its impact on *initialize* is diminishable as those for the trading direction (*buy*) and trading *quantity*. Nevertheless, those control variables do support the validity of Model I2, specifically the influences of risk aversion and emotion regulation strategies. Putting it all together, it is shown that on the one hand, high cognitive reflection leads to higher activity and – contrary to previous research – drives *liquidity taking*, as reappraisal does; on the other hand, risk aversion and suppression impels *liquidity providing*.

> **Result 1:** *Based on the analyzed personal attributes specific* trading behavior *can be identified.*

TABLE 5.3: *Trading Strategy*

Model	**I1** initialize	**I2** initialize	**I3** initialize
CRT_{high}	$.07^{***}$	$.09^{***}$	$.04^{**}$
	(5.20)	(6.23)	(2.81)
$TPL_{risk\ averse}$	$-.07^{***}$	$-.11^{***}$	$-.12^{***}$
	(−4.09)	(−6.02)	(−6.86)
$ERQ_{suppress}$		$-.09^{***}$	$-.07^{***}$
		(−5.95)	(−4.58)
$ERQ_{reappraise}$		$.16^{***}$	$.16^{***}$
		(10.57)	(10.22)
Control for products			✓
buy			.01
			(.61)
quantity			.00
			(−.89)
limit_price			$.00^{***}$
			(3.34)
(Intercept)	$.08^{***}$	.02	$.13^{***}$
	(5.27)	(1.19)	(4.49)
AIC	48,001.96	47,885.59	47,399.02
pseudo-R^2	6.42 %	6.74 %	8.11 %
Num. obs.	34,729	34,729	34,729

Notes: z-statistic in parenthesis; AIC: Sakamoto et al. (1986), pseudo-R^2: Nagelkerke (1991); $^{\cdot}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

5.4.3 Decision Quality

The final value of all stocks in the dataset is known. To answer the question *"What features does a trader need to be successful?"* the total profit of each trade is calculated based on the final value of the corresponding stock. Furthermore, each trading decision is classified according to its profitability (in other words: as 'right' or 'wrong'). From an ex-post perspective, the profitability (win) is considered as the *probability to make a profit*.

Trading Performance

In order to analyze the influence of a trader's personal attributes on her performance, a linear regression is conducted on the *profit* in P€ on a per order basis (Table 5.4). A strong

TABLE 5.4: *Trading Performance*

Model	**P1** profit	**P2** profit	**P3** profit
CRT_{high}	116.35^{***}	124.09^{***}	116.92^{***}
	(3.90)	(4.14)	(4.06)
$TPL_{risk\ averse}$	-160.99^{***}	-110.40^{**}	-91.64^{**}
	(−4.43)	(−2.95)	(−2.60)
$ERQ_{suppress}$		232.23^{***}	276.37^{***}
		(6.90)	(8.66)
$ERQ_{reappraise}$		8.30	-102.85^{***}
		(.26)	(−3.35)
Control for products			✓
buy			1794.55^{***}
			(64.18)
initialize			134.43^{***}
			(4.92)
(Intercept)	330.02^{***}	211.25^{***}	-816.32^{***}
	(9.75)	(5.17)	(−13.91)
Adj. R^2	.08 %	.23 %	11.44 %
Num. obs.	34,729	34,729	34,729

Notes: t-statistic in parenthesis; $^{\cdot}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

significance for cognitive reflection (standardized coefficients: .02 for P1 to P3) and risk aversion (std. coef.: −.02 for P1 and P2, −.01 for P3) can be seen in Model P1. The dataset shows that cognitive reflection, risk aversion, and usage of the suppression strategy (std. coef.: .04 in P2, .05 in P3) have significant influence on traders' performance (Model P2). In Model P3 even the reappraisal strategy has a significant influence (std. coef.: −.02). Nevertheless, the trading direction has the strongest effect on the traders' performance and the highest contribution to the profit (std. coef.: .33).

Probability to make a Profit

Comparing the OLS regression results for *profit* (Table 5.4, P3) with the logistic regression for *win* (Table 6, W3), only minor changes can be seen in the estimators' significance and direction. Interestingly, the traders' cognitive reflection ability has a significant influence in Model W3 *only* (Marginal Effect (mfx): .01). (Obviously, highly cognitive reflective traders do not robustly have a higher probability to make a profit, but *if* they gain, their average profits are higher.) The suppression strategy improves *decision quality* and slightly improves from W2 to W3 (mfx: .02 in W2 and W3). The usage of the reappraisal strategy had a higher significance level through the models – and additionally keeps its sign (mfx: −.01 in W2, −.02 in W3) as well. Contrary to that, risk aversion declines in significance and strength from model W1 to W2, but still beats a strong 5 % significance level in model W3 (mfx: −.02 in W1, −.01 in W2 and W3). Interestingly *initialize* plays no role in a trader's *probability to make a profit*. As earlier for *limit price* in Model I1 (Table 5.3), strong significances in combination with weak (marginal) effects for the control variables *quantity* and *limit price* can be seen in Model W3. Analogous to Model P3, *buy* has the strongest effect in Model W3 (mfx: .14).

Summing up, high cognitive reflection leads to better trading performance, whilst it does not (robustly) increase the probability to make a profit. Risk averse trader's performance is slightly worse, as are their chances to make a profit. Suppressors decide 'better' and are more likely to make a profit, whereas reappraisal tends to impair good decisions as well as the probability to make profits.

Result 2: *Personal attributes do significantly influence* trading performance *as well as the* probability to make a profit.

5.5 Conclusion

In this chapter, *advanced service analytics* was applied in order to gain comprehensive insights on participants' *market predisposition*. Based on a relatively short questionnaire, the trading history and regression models, it is possible to characterize participants' *trading behavior* and *decision quality* up to a certain degree. The applied methodology is hereby not tied to the context of play money prediction markets and can hence be used throughout similarly designed e-service systems like retail-trading systems. In particular, the influence

TABLE 5.5: *Probability to make a Profit*

Model	**W1** win	**W2** win	**W3** win
CRT_{high}	$-.02$	$-.02$	$.04^{**}$
	(-1.16)	(-1.43)	(2.65)
$TPL_{risk\ averse}$	$-.07^{***}$	$-.05^{**}$	$-.04^{*}$
	(-4.12)	(-2.64)	(-2.17)
$ERQ_{suppress}$		$.08^{***}$	$.09^{***}$
		(5.23)	(5.48)
$ERQ_{reappraise}$		$-.06^{***}$	$-.10^{***}$
		(-3.93)	(-6.51)
Control for products			✓
buy			$.58^{***}$
			(40.01)
quantity			$.00^{***}$
			(-7.31)
limit_price			$.00^{***}$
			(-3.54)
initialize			$.00$
			$(-.35)$
(Intercept)	$.34^{***}$	$.34^{***}$	$.04$
	(21.06)	(17.44)	(1.13)
AIC	46,493.89	46,464.39	44,060.04
pseudo-R^2	6.31 %	6.40 %	12.73 %
Num. obs.	34,729	34,729	34,729

Notes: z-statistic in parenthesis; AIC: Sakamoto et al. (1986), pseudo-R^2: Nagelkerke (1991); $^{\cdot}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

of a subjects' cognitive reflection ability, grade of risk aversion and use of emotion regulation strategies on trading behavior and decision quality in a play money prediction market was investigated. Putting all results together, it could be shown that cognitive reflection abilities have a significant positive influence on all investigated variables. One may argue that traders with higher cognitive reflection abilities performing better and having a higher probability to make a profit than the average is not very surprising. But traders in the high CRT group also behave differently: they submit more orders and tend to be liquidity takers. Interestingly, risk aversion has a positive impact on the number of submitted orders and a negative influence on a trader's performance as well as on her probability to make a profit. Finally, risk averse traders tend to be liquidity providers. Although neither emotion regulation strategy has a significant influence on a traders' activity, it can be shown that emotion regulation influences the initialization of trades: traders who confirm using the suppression strategy tend to provide liquidity, while the use of the reappraisal strategy leads to liquidity-taking trading behavior. Looking at the traders' performance, there is also a clear distinction between the reappraisal and the suppression strategy; traders who confirm using the suppression strategy make more profit on average and have a higher probability to make a profit, whereas traders who make use of the reappraisal strategy make less profit on average and have a smaller chance to decide profitable. Even if it may look like in these findings that the emotion regulation strategies reappraisal and suppression are opposite effects a person has to decide between, they are not. Even though both strategies seem to compensate each other in this study, one have to keep in mind that they are two strategies of emotion regulation a person makes use of 'simultaneously' in a different shape. Risk aversion has shown to affect the trading strategy towards liquidity providing. Furthermore, it slightly influences trading activity positively. In case of decision quality, risk aversion proved to be obstructive; both for profit and for the probability to make a profit.

Summing up, this study proved the possibility to categorize (potential) traders *ex ante* with *advanced service analytics*. The implications of these results are at least twofold. First, individual trading behavior can partly be predicted and therefore the market can be adapted accordingly. One possibility is to alter the user interface depending on the *market predisposition* of the particular user. A highly risk averse user with low cognitive reflection abilities who regulates his emotions mainly by using the suppression strategy is for instance less likely to need an order book since she tends to set limit orders instead of simply taking the quoted prices. Based on such knowledge, it is possible to create personalized and hence much clearer, user-centered trading interfaces. Second, a certain bonus/malus can be predicted a trader is going to experience in a market setting. This enables traders to self-assess their market predisposition and behave accordingly; e. g., by not joining a market. But even from the market providers' point of view, these results can be useful, since they can *ex ante* identify potential traders that do not have the 'right' predispositions. Additionally, they could identify potentially 'aptly traders' and recommend them to trade specific products. By following those implications, it should be possible to improve participants' decision performance within the context of prediction markets, which itself will lead to a better predictive power.

Chapter 6

Interpreting *Agent Behavior*: Reading a Trader's Mind

> "Thought is the seed of action; but action is as much its second form as thought is its first."
>
> RALPH W. EMERSON, SOCIETY AND SOLITUDE (1870)

6.1 Introduction

VOTERS, politicians, campaign advisers, businessmen as well as the media set out to consume and publish information in the context of an election. They all have a stake in the election outcome and seek to further their understanding of the election dynamic with up-to-date electoral probabilities. Public poll information, as a reflection of the public's take on the current political climate, helps campaign advisers measure their success and can influence informed voters' decisions. Before a major election, new polling results are published every other day in Germany by various institutes. As pointed out in Hillygus (2011), modern scientific polling has come a long way from its beginnings in 1937 and has seen an explosive growth in the last decades.

The Internet has democratized information and now plays a crucial part in every election. In fact, a recent study by Bitkom (2013) shows every third German to consider the

Internet the decisive factor in the upcoming election.[1] One of the most promising ways of forecasting elections are prediction markets.

In prediction markets, participants trade contracts whose payoff depends on the outcome of uncertain future events. For example, a market contract might reward a dollar if a particular presidential candidate is elected. An individual who thinks the candidate has a 65 % chance of being elected should be willing to pay up to 65 cents for such a contract. Market participants form expectations about the outcome of an event. Comparable to financial markets, they buy, if they find that prices underestimate the probability of the event in question and they sell a stock, if they find that prices overestimate the probability of an event. The *Iowa Electronic Markets*[2] (IEM) are a well-known example for prediction markets. One of the IEMs' markets, the *Iowa Presidential Stock Market* (IPSM), is a political stock market (PSM) which predicts, inter alia, the outcome of U.S. presidential elections Forsythe et al. (1992). The IPSM features contracts that represent one nominee each. Market participants buy and sell nominee contracts depending on their assessment of the U.S. presidential election outcome.

PSMs have been used widely in different countries and electoral systems (e. g., Berg et al., 2008; Forsythe et al., 1992). In contrast to the traditional, straight-forward process of a representative part of the population eligible to vote answering a question like *"What party would you vote for, should the election take place this Sunday?"*, in a PSM participants are incentivized to trade on their *expectation* about the election outcome. Hence there are two distinct differences. First, participants provide their *beliefs* about the election outcome, opposed to simply stating his political *preference* in a poll. Second, the market mechanism incentives early, timely and accurate predictions about the outcome.

One thought that immediately comes to mind is that traders may be biased by their personal preferences in their valuation of contracts. This would be no surprise, dealing with a sensitive topic like politics. The judgment bias is well known from sports gambling, where devoted fans of sports clubs show a substantial amount of wishful thinking Babad and Katz (1991). In this study, the way political preferences help to shape traders' decisions in a German PSM is analyzed. It is found that traders excessively buy the party they prefer to win the election. The bias differs in strength for different parties but is nonetheless consistent, no matter what party preference. In general, it seems that it is most pronounced for small parties. This result even holds for the small but identifiable subgroup of tactical

[1]The study relates to the German federal elections on 2013-09-22.
[2]Accessible at the URL `http://tippie.uiowa.edu/iem/` .

voters. As this bias is so consistent for subgroups and preference, a straightforward prediction model can be provided to infer traders' party preferences by analyzing the trading behavior. This might reduce participants' *perceived* as well as their *effective anonymity* in prediction market and thus their behavior.

The remainder of the chapter is organized as follows: First, a closer look at the core ideas behind PSMs is taken to explain the way they work and why they have been found to be so successful. A review of the literature regarding bias in prediction markets is also provided, especially in PSMs and in light of the research questions. Subsection 6.3.1 provides specifics about the market on which this research is conducted. The empirical results of this research are discussed in Section 6.4. Finally, Section 6.5 concludes.

6.2 Related Work

6.2.1 Political Stock Markets and Polls

Political Stock Markets – as a subset of prediction markets – share their main objective, namely aggregating information from its participants in order to create efficient real-time forecasts for uncertain future events. In this case, these uncertain future events are of political nature, i. e., elections, nominations for elections or policies.

As a forecasting method prediction markets offer many advantages. First off they provide the incentives for traders to truthfully disclose their information and an algorithm to weight opinions (Arrow et al., 2008). Compared to statistical forecasting methods, these markets can incorporate real-time information. As prediction market prices are updated immediately when traders incorporate their expectation in prices, they provide continuously and timely updated forecasts. Compared to eliciting expert opinions, prediction markets eliminate the effort of identifying experts and motivate their participation. In most cases they allow anonymous participation, which may increase the likelihood of nonconformists to participate and reveal information and they do not need to deal with conflicting opinions.

The question of PSMs' performance compared to polls has sparked some attention in the last years. Berg et al. (2008) analyze the results of more than ten years' worth of PSM predictions on the IEM against corresponding polls and conclude that market results outperformed the polls in most cases. Similarly, Berlemann and Schmidt (2001) find that

– though by a less broad margin – European PSMs significantly outperformed respective polls as well. There has been some doubt in respect to the naive manner polls were used in their comparisons, i. e., Erikson and Wlezien (2008) argue that polls needed to be properly adjusted before comparison, but as Rothschild (2009) points out, fairly adjusting both PSM and poll results yields PSM as the overall most accurate predictor again.

6.2.2 Biases in Political Stock Markets

In theory, there exists an ideal called *rational trader*. He is always instantly available to trade when an opportunity arises to make a profit, maximizing which is his only objective. If he is exposed to new information, he objectively incorporates it and updates his beliefs accordingly. As all too often, reality tells a different tale: Traders act imperfectly out of a variety of reasons, and PSMs are no exception from this rule. Different types of bias, mostly already known from regular markets, betting markets, polling and other fields, have been identified in various studies. A general consensus seems to be that traders' judgment of probabilities can be impaired by favorite-longshot (cf. Wolfers and Zitzewitz, 2004; Snowberg and Wolfers, 2010) and *judgment bias*, while it is unclear whether these individual biases sway the market on an aggregate scale. Here the focus is on the judgment bias.

Anyone who has recently discussed the odds of a sports event with supporters of both teams is very likely to know this effect. Supporters generally tend to overvalue their team and therefore experience a judgment bias when predicting the outcome. Sports enthusiasts show a significant amount of this aforementioned *wishful thinking* even after explicitly being asked to stay objective (Babad and Katz, 1991).

Multiple subsequent studies (i. e., Babad et al. (1992); Babad and Yacobos (1993) as well as Uhlaner and Grofman (1986)) show that a considerable amount of wishful thinking is also observable when it comes to politics. While the amount of wishful thinking in a sports context is dependent mostly on emotionalism and level of fanhood, in case of politics the preferred party plays an important role. The intensity of wishful thinking decreases, moving from right-wing towards left-wing on the political spectrum. Interestingly, extreme left-wing supporters even show an inverse effect. Since the aforementioned studies measure voters' intentions by inquiring predictions about election outcomes, it seems natural that a *judgment bias* caused by wishful thinking is present in PSMs as well.

As presumed, nearly all authors who investigate this effect in their PSM experiments, report significant amounts thereof. Take Forsythe et al. (1992), who find the judgment bias affecting trading behavior on average and most traders incapable of valuing prices objectively. These results are replicated by Forsythe et al. (1998), Jacobsen et al. (2000), and Berlemann and Schmidt (2001). Forsythe et al. (1999) provide a detailed discussion about judgment bias and two effects that can cause it: (i) the *false consensus effect*, which states that traders overestimate their own representativeness, and (ii) the *assimilation-contrast effect*, which describes the tendency to interpret information overly in the direction of one's own preference. They find that although most individual traders are significantly biased in their trading, overall market prices are not. As an explanation, a fraction of traders is assumed to be aware of biased traders' shortcomings, correcting market prices while taking advantage of this information. Since they tend to set limit orders close to market prices, they are known as marginal traders. However, this result does not seem to hold universally: Berlemann and Schmidt (2001) find a judgment bias on the aggregate scale in German PSMs.

6.3 Setting and Research Questions

6.3.1 Experimental Setting

A German PSM is used to examine the effects of traders' political preferences on trading activity . Specifically, data from the PIX for the German federal election 2013 is used. For a detailed description of the PIX market refer to Subsection 4.3.1.

6.3.2 Research Questions

As the wishful thinking bias is so persistent, two questions arise. First, *are all subgroups equally biased in their trading decisions?* One hypothesis is that different party preferences lead to a more pronounced bias. Second, *are tactical voters as biased as preference based voters?* Finally, if the bias is stable and constant over subgroups, research question 3 as presented in Section 1.2 arises:

> **Research Question 3:** *How well can an unobtrusive analysis of trading behavior reveal trader preferences?*

6.4 Results

In this section the empirical findings are presented, starting with descriptive statistics. Subsequently, traders' self-assessed party preferences are reported. Finally, traders' party preferences are predicted.

6.4.1 Descriptive Statistics and Methodology

A day before the election, the average transaction prices were collected for all parties and a final prediction was created. Table 6.1 illustrates the market prediction and displays the election outcome for comparison. In general there was a lot of uncertainty about the election outcome, due to potential strategic voting.

TABLE 6.1: *Market Prediction and Election Outcome*

	CDU/CSU	SPD	LINKE	Grüne	FDP	AfD	Piraten	Rest-of-field
Prediction	35.06 %	20.32 %	8.47 %	7.63 %	6.70 %	15.15 %	2.53 %	4.15 %
Outcome	41.55 %	25.74 %	8.59 %	8.44 %	4.76 %	4.70 %	2.19 %	4.03 %
Abs. Dif.	6.49 %	5.42 %	.12 %	.81 %	1.94 %	10.45 %	.34 %	.12 %

Measuring the Judgment Bias

The wishful thinking judgment bias has already been covered in Subsection 6.2.2. Using the questionnaire functionality, users willing to share this information were matched to their preferred party. This enables to analyze the extent of individual false consensus effect in the spirit of Forsythe et al. (1992), as a proxy of judgment bias. From 2013-06-21 until 2013-07-31 and from 2013-08-26 until 2013-09-23, the question *"Which party can you identify with the most, when it comes to national politics?"* (cf. Sjöberg, 2009) was run. Possible answers were *CDU/CSU*, *SPD*, *FDP*, *Grüne*, *DIE LINKE*, *Piraten*, *AfD*, *another party* and *prefer not to say*. Only one selection is permitted, allowing to match each participating trader to exactly one party. For the subsequent analysis, it is assumed that this question is truthfully answered and that preferences are valid and remain constant for the runtime of the two markets.

The analysis of the individual false consensus effect is approached with the known methods (cf. Forsythe et al., 1998). One possibility to determine whether supporters of a party systematically preferred the corresponding contract is to analyze their portfolios.

The judgment bias states that the preferred contract is overvalued, which should cause a higher demand for that contract from affected traders than from the average trader. The logical conclusion is that since he is expected to invest more money, the value share of the preferred contract in a biased trader's portfolio exceeds the average trader's value share in this contract. Specifically, the following measure is defined:

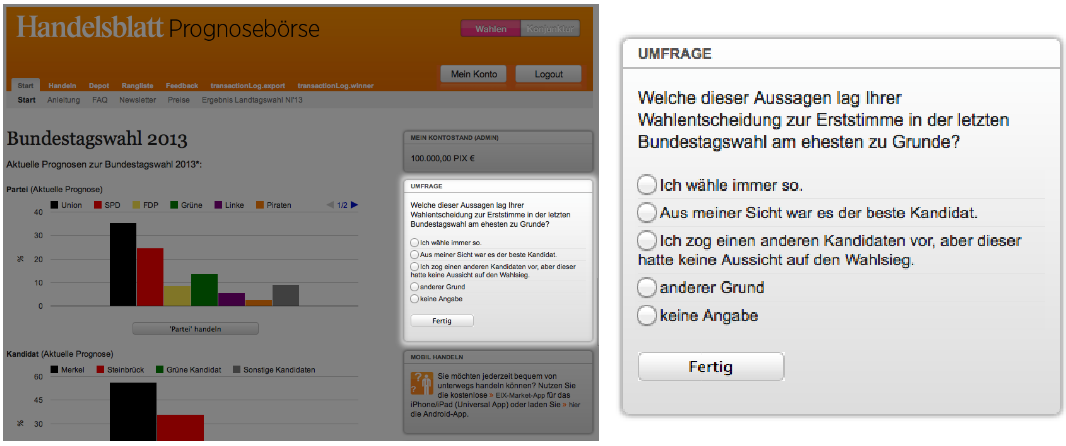
$$\xi_{t,party} := \frac{v^i_{t,party}}{v^i_t} \frac{\sum_{i=1}^m v^i_t}{\sum_{i=1}^m v^i_{t,party}}, \tag{6.1}$$

where $i \in \{1, \dots, m\}$ is the position of party (with the captions *CDU/CSU*, *SPD*, ..., *rest-of-field*; leading to $m = 8$) in the portfolio. Vector v_t is obtained by multiplying the 'unbalanced portfolios' (Forsythe et al., 1999, p.92) with the market prices at t aggregated for all users that responded to the party preference questionnaire. Whereas, $v_{t,party}$ corresponds to the multiplication of the 'unbalanced portfolios' with the market prices at t aggregated for all users that prefer *party*. Hence, $\xi_{t,party}$ describes the ratio of party's value percentage in the supporters' portfolio compared to its value percentage in the aggregate portfolio. Bearing the prior assumptions in mind, an existing, notable judgment bias would yield $\xi_{t,party} > 1$. This is precisely what is empirically calculated in the portfolio analysis for judgment bias.

6.4.2 Traders' Reported Preferences

The PIX was used in several ways to collect data for this study. Since the author co-designed the market and the underlying data structure, it was possible to store literally every piece of information needed on events that take place on the PIX. Since the author co-designed the market and the underlying data structure, it was possible to store all information necessary to conduct this study. First and foremost, detailed information on events such as orders and transactions was used.

This includes, e. g., whose order was matched on which exact date, and the price and volume that was subsequently traded. Second, a survey function was added to the platform: The questionnaire (Figure 6.1) allows for questions with a predefined list of answers to appear on the PIX's main page, one at a time. Adjustable settings include the question mode, allowing only one reply selection or multiple. Equally important, it can be precisely defined when, where and in what order questions are asked. These properties are defined by a time window for question activity, and a priority ranking for the order of appearance.



(A) *Market-embedded Survey* (B) *Magnification*

FIGURE 6.1: *Screenshot of the Market-embedded Survey Functionality (a: placement on website, b: magnification)*

Obviously, it was made sure that every user can answer every question only once. For each of the questions, the option "prefer not to say" was provided, due to the sensitive nature of the questions. Users submit their responses by selecting the radio button or checkbox and clicking submit. A self-made questionnaire infrastructure was used because this way, traders can quickly reply to a question or two without leaving the website. Hence, willing responder who do not want to go to an external website were not lost. 336 traders participated in the experiment on the false consensus effect by providing their personal political preference. Table 6.2 lists the results. Using this data, it is possible to measure for judgment bias.

TABLE 6.2: *Party Preference*

	CDU/CSU	SPD	LINKE	Grüne	FDP	AfD	Piraten	Rest-of-field
Answers	28	25	16	22	23	204	8	10
Percent	8.33 %	7.44 %	4.76 %	6.55 %	6.85 %	60.71 %	2.38 %	2.98 %

Notes: N = 336

Using the questionnaire function, a subset of traders was also identified that can be regarded as tactical voters. According to a measure developed by Heath et al. (1991), called 'Heath et al. measure' in Fisher (2004), tactical voters can be directly identified using a simple question. The Heath et al. measure was slightly modified for the context of the German situation. Table B.2 in Appendix B depicts the full measure. This study set out anticipating that tactical voters exhibit a difference in behavior, also on a PSM. The results of the questionnaire are that 21.6 % of traders belong to the class of tactical voters based on their answers.

6.4.3 Traders' Predicted Preferences

In this section, the empirical results from the collected data on the PIX including 2013-09-22 are presented. First it is shown that participants exhibit a strong and significant judgment bias on the PIX. This result corresponds to previous findings in the existing literature on judgment bias in prediction markets.

Using the party preference data as presented in Table 6.1 and traders' portfolio on the test day $t = 2013/09/22$, the judgment bias is measured. The results for $\xi_{t,party}$ are listed in Table 6.3. Recall that a $\xi > 1$ means that a group of traders holds a greater value in their own preferred contract than all traders on average. The result indicates that this is true for all major parties. For supporters of CDU/CSU, who 'only' hold 62 % more in their own contract, the effect is not very strong. Parties with small numbers of participating supporters, such as LINKE and Piraten, seem to be the ones to rely the most on their own preferred contract. However it must be kept in mind that there is no statement about significance of the effect yet. A smaller number of traders allows for higher variance. (The rest-of-field contract has been left out since the concept of 'rest-of-field-supporters' does not make sense.)

TABLE 6.3: *Judgment Bias per Party*

	CDU/CSU	**SPD**	**LINKE**	**Grüne**	**FDP**	**AfD**	**Piraten**
$\xi_{t,party}$	1.62	2.05	7.23	3.29	4.00	1.95	9.30
N(*party*)	27	23	14	20	21	153	6

Notes: N = 264

This study aims to examine, if an individual's preference for a given party is linked with his 'biased investment characteristics'. As biased investment characteristics, the ratio of an individual's investment in stocks of that particular party to the individual's overall investments is used. Hence, one simple OLS regression is conducted per *party*; describing the 'investment characteristics' by an intercept and a dummy variable indicating the individual's preference for *party* (Table 6.3). All seven regressions are significant to the 0.1 %-level.

As a robustness check a time series with the values of ξ is constructed, since they are time dependent. In order to do this, all relevant data (like portfolio structure) must be counted back in time. In general ξ is well above 1 for all parties and increasing over time towards the election. The time series of ξ values provide a good robustness check. From the data it can be seen that the judgment bias is aggravated for the underdog parties[3] (FDP, Grüne, LINKE, Piraten, AfD) compared to the established parties (average of 5.15 vs. 1.84).

According to the adjusted Heath et al. questionnaire 21.6 % of answerers are identified as tactical voters which seems to be a relatively high figure. Unfortunately there is no reference number for Germany as a whole. One might assume that tactical voters do not exhibit the wishful thinking bias to the same extent as non-tactical voters. These two classes of participants are compared and no statistical difference can be found (average ratio invested in preferred contract: .43 vs. .39, p-value: .72).

Finally, a model was built to predict party preference by analyzing portfolio data. Specifically, a simple tree based classification model is used. The party preference of a participant p is predicted using the percentage of invested play-money per party and additionally the net number of shares bought in each party.

In order to test the validity of the model the sample ($N = 264$) is split in a training (~62 %; 164 observations) and a validation (~38 %; 100 observations) set. In the out of sample test, the model correctly classifies 70 % of all instances. The party wise detection rates (e. g., AfD: 93 % vs. CDU/CSU: 53 %) suggest that a higher number of supporters in the training set leads to better results. Keep in mind that the base rate is one in eight or 12.5 % – given that each participant could prefer one of the 8 parties. Moreover, as the most intuitive and straightforward method is used to model the data, more predictor variables (e. g., gender, age, other trading behavior or activity) or better methods such as random forest, support vector machines (SVM), or neural networks are very likely to yield better models. Hence, it can be concluded that prediction market data enables researchers and practitioners to classify their trading population very easily.

[3] Here, a party is classified as underdog, when their election outcome is below 10 %.

6.5 Conclusion

The motivation behind this research is to understand how the increasing importance and possibilities of online prediction models change the way people think of events like elections. While elections used to be like 'blackboxes' in past times, they seem to have become predictable, almost to the point where the candidate with the best data analysts will win the election such as stories from Issenberg (2012) would have made us believe.

Political Stock Markets are one of the new continuous available prediction methods. They are based on the assumptions that market prices (in an abstract form of predictions) are set by rational unbiased traders. The key question that is addressed in this study is whether and to which extent traders stay objective or if they are biased by their own preferences.

The participants' preferences are collected using questionnaires directly accessible from the trading website which are simple and do not require much time. This helps obtaining a quite high number of answered questions and subsequent data to analyze (cf. Chapter 9). Surprisingly, although this is the most personal question asked amongst multiple other questions, it is still the most frequently answered question. It seems like traders are very eager to identify themselves as supporters of their preferred parties, which leads to the conclusion that most traders have strong opinions and that their political opinions are among the reasons for trading.

This strong opinion does influence how they trade and act in the market, even though they are incentivized to not do. Through a portfolio level analysis of trading data matched with survey data it is possible to consistently predict voter intention in the market population. Moreover, evidence is provided that the bias is consistent over all parties but elevated for underdog parties. Surprisingly, analyzing subgroups no difference is found in the bias between tactical and non-tactical voters.

As the bias is so consistent for subgroups and preference, it is possible to provide a straightforward prediction model to infer a trader's party preference by analyzing his trading behavior with 70 % accuracy. This is important because it might reduce participants' *perceived* as well as their *effective anonymity* which is sometimes highlighted as a major reason for prediction markets' success.

Chapter 7

Extending the *(IT) Infrastructure* into the Mobile World: Comparing Trading Performance in Stationary and Mobile Settings

> "I have always wished for my computer to be as easy to use as my telephone; my wish has come true because I can no longer figure out how to use my telephone."
>
> BJARNE STROUSTRUP, 1990

7.1 Introduction

WE often rely on information systems (IS) to filter, aggregate, and present information we need in a manner that supports decision making processes. A common misbelief about decision making is that the more information available, the better our decisions. In contrast, it has been shown that more information can lead to decreased decision making performance, e. g., due to *information overload* (cf. Malhotra, 1982). For electronic markets, Teschner et al. (2011, 2014) showed that more information can be harmful for individual trading performance.

With the rise of mobile information systems, the question arises how decision behavior and decision performance are influenced by its usage. In general, two major developments seems to influence the usage of IS here. First and obviously, mobile IS enables the usage of IS in a mobile context. Hence, one is enabled to make use of IS in settings where it was not possible before. This opens a whole new set of opportunities (e. g., Muntermann and Janssen, 2005) and hindrances which can, for instance, result in faster reaction times to news in a trading context as well as to a higher degree of distraction or uncertainty. Second, mobile IS might also supersede traditional IS in certain settings. Mobile Human-Computer Interaction (HCI) often differs from its stationary counterparts (e. g., Schmiedl et al., 2009) for a variety of reasons: inter alia different screen sizes (e. g., Brewster, 2002; Adipat and Zhang, 2005), gesture controlled vs. mouse and keyboard, operation systems, and reliability of network connectivity. Thus one might expect that these differences do result in different outcomes in some cases. Therefore it is expected that users of mobile IS will perform differently than users of stationary IS for a given task or problem.

In order to design mobile systems that support good decision making it is necessary to analyze *how participants search for information* and *how they incorporate this information in their decision process*. Moreover, *behavioral aspects of IS users have to be linked with the quality of their decisions* in order to improve the design of mobile IS. More precisely, this study tries to answer the following question: *"How do different devices (and therefore user interfaces) affect decision behavior and decision outcome?"* Hence, a field study is conducted on an electronic market to shed some light onto this higher research question.

Specifically, the research is conducted in a repeated market environment called *Kurspiloten* (cf. Section 4.2). The Kurspiloten market is a prediction market (cf. Wolfers and Zitzewitz, 2006; Luckner, 2008) designed to forecast the stock exchange value of selected stock indices and commodities on a weekly basis. This prediction market is set up as a continuous double auction, like in financial markets, with one stock representing each new release of economic information. Participants buy if they think that prices underestimate the probability of an event and sell if they think prices overestimate the probability of an event. The prediction market thereby aggregates information in the same way a stock market does, which is relatively efficient in an ex-ante information sense. In the Kurspiloten field experiment with nearly 2,000 active participants the impact of mobile and stationary interfaces on user behavior and decision performance is studied.

The remainder of the chapter is organized as follows: Section 7.2 discusses related literature on decision making in the context of stationary and mobile information systems. The experimental setting, methodology used, and research direction is presented in Section 7.3. Subsequently, the results are discussed in Section 7.4. Finally, Section 7.5 concludes this chapter.

7.2 Related Work

7.2.1 Information Systems and Participant Decisions

Kauffman and Diamond (1990) highlight the importance of research on behavioral decision making and information presentation effects. They examine how behavioral effects may become operative in screen-based securities and foreign exchange trading activities, where users can choose among information presentation formats that support trader decision making. They present a model to identify where and how information, heuristics, and biases might affect decision making in trading environments. In the domains of decision support systems and online shopping environments the influence of the interface on decision behavior has been repeatedly demonstrated. Kleinmuntz and Schkade (1993) find that information displays do influence decision processes by facilitating some decision strategies while hindering others. Decision makers balance the desire to maximize accuracy against the desire to minimize effort. They further separate characteristics of information displays into the form of individual items (numerical, verbal or pictorial), the organization into meaningful structures (groups, hierarchies or patterns) and the sequence (the order in which information element appears). In a follow-up study they show that organization strongly influences information acquisition while form influences information combination and evaluation. Sequence had only a limited effect on information acquisition (Schkade and Kleinmuntz, 1994). Investigating the relationship between problem representation and task type in information acquisition, Vessey and Galletta (1991) develops the cognitive fit theory. The theory proposes that the correspondence between task and information presentation leads to superior task performance for individual users. In several studies, cognitive fit theory has provided an explanation for performance differences among users across different presentation formats such as tables, graphs, and schematic faces (Vessey and Galletta, 1991; Vessey, 1994). Additionally they show that increasing interface flexibility instead of an informed choice of display format may be harmful rather than helpful to the problem solver. Similarly Speier and Morris (2003) compare the use of visual and text-based interfaces for low and high complexity tasks. They find that in low complexity environments participants perform better using text-based query tools.

However in high complexity environments participants perform better with visual support. Turning to the optimal pool of available information in decision support systems, empirical work has shown that users can handle only a certain amount of data.

Malhotra (1982) concludes that individuals cannot optimally handle more than ten information items or attributes simultaneously. Testing decision accuracy, Streufert et al. (1967) show that as information load increases, decision making first increases, reaches an optimum (information load ten) and then decreases. Finally, in an interactive home shopping simulation, Ariely (2000) tested how the participants' control over information influences their utilization of this information. He compared four settings: if information control was high-low and the task complexity was low-high. He finds out that increased control over information leads to better performance in tasks with low complexity and lower performance in the high complexity setting. He reasons that participants in the low complexity setting, when demand on processing resources is low, more information is beneficial. However, in complex situations the information is detrimental to performance due to the additional burden of selecting the right information (Ariely, 2000). He concludes that when cognitive load is high (e. g., when the task is novel or difficult) high information control can be harmful.

To summarize previous work, the amount and control of information, as well as the information representation does influence user behavior. On the one hand information control improves performance by improving the fit between actions and outcomes. On the other hand information control requires the user to invest processing resources in managing the information amount and flow. As a conclusion, information control has both positive and negative effects on performance. The two tasks of processing and managing information are related and codependent. Finally, one must note that previous work has mainly investigated the topic in laboratory settings. User behavior and decision performance is analyzed in a field experiment setting, namely a prediction market.

7.2.2 Comparing Stationary vs. Mobile

Eriksson (2012b) compares the online self-arrangement experience of mobile device users to stationary computer users in an electronic travel service experiment. Thereby he focuses on the three dimensions efficiency, effort, and anxiety. He found that mobile device users experiences the given task more negative. In a follow-up study Eriksson (2012a) compares the use of different channels in electronic travel services in Finland for 2004 and 2011. There he found that most customers use only a computer to fulfill travel related tasks. Nevertheless, there are a small and growing number of customers using only mobile devices for such tasks. Interestingly, the number of customers using both computers and mobile devices is much larger than the number of (solely) mobile device users. Altogether, most users report the computer as their preferred interaction channel. Both studies focus on user perception of the offered services.

In contrast to the studies mentioned earlier, Muntermann and Janssen (2005) focuses on behavior and outcome depending on the used channel. They find that mobile financial information systems can provide serious benefit to customers' value. They investigate realizable returns in a stock market subject to the users' reaction time to incoming events. In a simulation, based on a real-world dataset, they compare two different scenarios of information latency. The results show that, in the low-latency scenario, customers gain more than five percent of realizable returns compared to less than 2.5 percent in the high-latency scenario. Due to the characteristics of mobile information systems users are able to react nearly immediately to new information and transform their advantage into monetary gains. In a market-based environment, Teschner et al. (2012) describe a method to distinguish between decision supporting and misleading information in mobile applications. Their preliminary results suggest that the decision making process differs depending on the device used. Besides the studies mentioned there exists hardly empirical work analyzing decision performance in mobile applications.

7.3 Setting and Research Questions

7.3.1 Experimental Setting

This study was conducted on the Kurspiloten market. Kurspiloten is a prediction market for selected stock market indices and commodities. Subsection 4.2 contains a detailed market description.

Operationalization

In the following, the indicator variable $Device_o$ is 1 for a given order o when it was submitted trough the mobile application, otherwise it is 0. The order type used for a given order o is described by the indicator variable $Market\ Order_o$. It is 1 for a liquidity taking market order and 0 for a liquidity providing limit order that cannot immediately be matched. For a buy order o the dummy variable TD_o ("trading direction") is 1, whereas for a sell order it is 0. In this continuous market the outcome of each stock (i. e., the final value) can be observed. Therefore the information content of each order can be measured ex-post. With respect to the outcome of a stock, if the order moved the price in the correct direction it is classified as informed, whereas an order moving the price in the opposite direction of the outcome is classified as uninformed. Based on Teschner et al. (2011), the following score is used to capture this process:

$$Score_{o,i} = \begin{cases} 1, & price_{o_i} \leq f_{v_i} \quad \text{and} \quad o_{type} = BUY \\ 1, & price_{o_i} \geq f_{v_i} \quad \text{and} \quad o_{type} = SELL \\ 0, & price_{o_i} > f_{v_i} \quad \text{and} \quad o_{type} = BUY \\ 0, & price_{o_i} < f_{v_i} \quad \text{and} \quad o_{type} = SELL \end{cases} \tag{7.1}$$

The price of an order o for the stock i is represented as $price_{o,i}$. The fundamental final outcome value of a stock is represented by f_{v_i}. In a way, the $Score_{o,i}$ can be interpreted as an indicator for the profitability of an order and thus as the decision outcome of a trader; $Score_{o,i}$ is 1 for a profit greater equals to zero and 0 otherwise.

Analyzing Decision Confidence

As described in the last section, two proxies are used to measure the participants' decision confidence and trading behavior. The quantity of a specific order is related to the device used. As the different stocks exhibit different historic variances (e. g., the Dow Jones is much more volatile than Bund-futures) the analysis is controlled for these variances by adding the market indicator variables M_i. These control variables are included in all regression models. To identify the influence of the device on the submitted quantity (first confidence proxy) the following OLS regression is used:

$$Quantity_o = i + \beta_1 \times Device_o + \beta_2 \times Market\ Order_o + \beta_3 \times TD_o + \sum_{i=1}^{12}(\gamma_i \times M_i) \tag{7.2}$$

For the second proxy it is necessary to look at *how* users submit their orders. For an executed order there are only two possibilities; either an order is a market order (price taking) or a limit order (liquidity providing). The market order *initializes* a trade by getting immediately matched against a standing limit order and thus taking liquidity from the market. On the contrary, orders are counted as limit orders when they cannot immediately be matched and thus executed. In this case, they are written to the order book and hence provide liquidity to the market. As this is a binary outcome, a binomial logistic regression model is used. If an order is initializing a trade, the dependent variable is 1 otherwise it is 0. Equation 7.3 measures the device's influence on the probability whether an order is a market order (i. e., price taking) or a limit order (i. e., liquidity providing).

$$log\left(\frac{\pi_{Market\ Order}}{\pi_{Trade}}\right) = i + \beta_1 \times Device_o + \beta_2 \times Quantity_o + \beta_3 \times TD_o + \sum_{i=1}^{12}(\gamma_i \times M_i) \tag{7.3}$$

Analyzing Trader Performance

In order to calculate the influence of trader behavior on trading outcome Equation 7.3 is adapted the following way. The dependent variable is the score (profitability) as defined in Equation 7.1, which is 1 for a profit and 0 for a loss. As before, the analysis is controlled for different risks in the market categories by adding the dummy variables M1-M12 and receive the following equation:

$$\begin{aligned} log\left(\tfrac{\pi_{Score}}{\pi_{Trade}}\right) = i \quad & +\beta_1 \times Device_o + \beta_2 \times Market\ Order_o + \beta_3 \times Quantity_o \\ & +\beta_4 \times TD_o + \textstyle\sum_{i=1}^{12}(\gamma_i \times M_i) \end{aligned} \tag{7.4}$$

The regression for the profit (Equation 7.5) is analogous to Equation 7.4, except that the quantity has to be dropped due to the obvious high correlation with profit:

$$Profit_o = i + \beta_1 \times Device_o + \beta_2 \times Market\ Order_o + \beta_3 \times TD_o + \sum_{i=1}^{12}(\gamma_i \times M_i) \tag{7.5}$$

7.3.2 Research Questions

As more decisions are facilitated through mobile decision support systems, one of the most urgent questions is *"How to design interfaces that improve decision making?"* In order to answer this higher research question it has to be deeply understood if and how the interface influences decision making. More specifically it needs to be analyzed how participants search for information and how they incorporate this information in their decision process.

The second goal of this study is to link behavioral aspects of the market participants with the quality of their decisions. Creating a link between behavioral aspects of the participants and quality is important since the quality of the predictive power is directly negatively affected if participants make systematically biased decisions. Hence, the second research question is *"How do different devices (and therefore user interfaces) affect decision behavior and decision outcome?"* The main research question of this study is introduced in Section 1.2 as research question 4 and reads:

Research Question 4: *Are decision behavior and decision outcome affected by the kind of device used?*

The experimental set up is well suited to studying the behavioral aspects of decision making because in contrast to financial markets (i) the outcome of events in the market is ultimately known and (ii) the ex-post trading performance of participants can be measured. To give indications for these research questions, the influence participants' device choices have on their information usage is analyzed. Following Ariely (2000), it is expected that users choose different information items in order to adapt the interface to their informational needs, needed to fulfill the given (decision) task (i. e., trade). Due to mobile devices' limitations, it is expected that this adaption leads to less consumption of information items in case of mobile device users; whereas users of a stationary device consume (slightly) more information items as they do not suffer from device related limitations.

Following the research model (Figure 7.1), based upon van Witteloostuijn and Muehlfeld (2008), the *Device* usage is connected with the participants' *Confidence and Trading Behavior* in hypotheses H1. In order to measure the vague concept of *confidence*, two common proxies are used: a) the *order type* used to trade and b) the *order size* in stocks. Particular, the *order type* is distinguished between *market orders* and *limit orders*. A *limit order* is executed for a given *limit* price, whereas a *market order* is an order that executes immediately, but often at less favorable prices. Hence, the trader submitting a *market order* pays the spread (i. e., the difference between the best buy and best sell prices), but is guaranteed immediate execution. A less confident trader is less willing to trade immediately and willing to wait for the market to move in his direction (Teschner and Weinhardt, 2012). The second proxy to quantify trader behavior in financial markets is the *size* of the submitted order (cf. Yang et al., 2012). As all traders have the same start portfolio the quantity of a trade is a proxy for a trader's confidence perception. If a trader has doubts about the future development of an indicator he is likely unwilling to bet all on one – or just a few – shot(s).

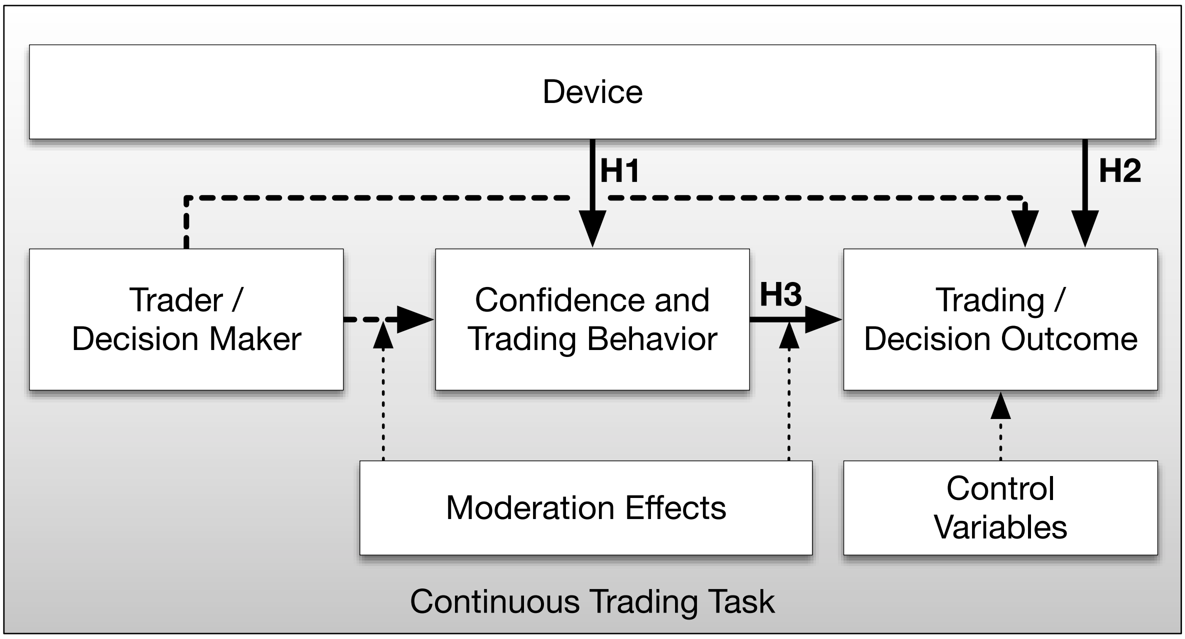


FIGURE 7.1: *Research Model*

Summing up, the users' *device choice* is connected with their decision *confidence* and *trading behavior* by using the proxies *order size* and *order type* (market order or limit order). Hence, the following hypotheses are stated:

Hypothesis 1a: *Participants using mobile device submit orders with lower average quantity.*

Hypothesis 1b: *Participants using mobile device have a lower probability to submit market orders.*

In the next step, the interplay between *Confidence and Trading Behavior* and *Trading/Decision Outcome* (H2 in Figure 7.1) is regarded. As described above, large orders are expected to be more informative than smaller orders. In other words, an increased order-size is expected to be a predictor for a profit (Hypothesis 2a). Previous research on trading behavior showed that traders who set prices and use limit orders (market-making) are less mistake-prone and appear to be more rational than traders using market orders (price-taking) (Oliven and Rietz, 2004). Hence, traders using market-making trades (limit orders) are expected to be more successful (Hypothesis 2b). The hypotheses related to H2 (Figure 7.1) are as follows:

Hypothesis 2a: *Increased order-size is positively correlated with the resulting profit.*

Hypothesis 2b: *Participants using limit orders are more likely to submit profitable orders.*

Finally, and most importantly, by controlling for *trading behavior* it is analyzed how the self-chosen *Device* influences the participants *Trading/Decision Outcome* (see H3 in Figure 7.1). The *decision outcome* of a submitted order can be analyzed depending on the resulting profit or loss. The needed heuristic is detailed in Subsection 7.3.1. The intuitive reasoning is that more information can be displayed, understood, and incorporated using the stationary device. As more information are expected to be beneficial a better trading performance is expected (Hypotheses 3a and 3b). Moreover, one could argue that participants can trade on ad-hoc information just when they are available and may use this advantage to make a profit (cf. Muntermann and Janssen, 2005). In order to proxy *decision outcome* two measures are used. The first is the likelihood that a trader makes the 'right' decision (i. e., she submits a profitable order; for details see Equation 7.1). Second, the profit resulting from each order can be measured ex post. Thus the hypotheses for the interface influence on decision accuracy (see H3 in Figure 7.1) are:

Hypothesis 3a: *Participants using mobile device are less likely to submit profitable orders.*

Hypothesis 3b: *Participants using mobile device make less profit.*

Those three steps combined provide an indication of the market interface and the informational impact on trader behavior. Moreover, they provide insight into the inter-play between device, information and decision making.

7.4 Results

This section summarizes the results of the conducted study. The three hypotheses presented in Section 7.3.2 are analyzed following the research model and the results are interpreted. First, the device influence on traders' decision confidence is reported. Second, trading behavior and trading performance is analyzed.

7.4.1 Decision Confidence

Users accessing the platform through the mobile device were expected to use on average small order sizes (Hypothesis 1a). Using a simple t-test no significant difference was found between the order-sizes (*quantity*) of web and mobile traders (web: 1,038.28; mobile: 1,276.55; t-stat = .84). Even when controlling for different factors (as described above), no significant influence of the device used was found (Table 7.1). Turning to what *type* of orders (i. e., market order vs. limit order) participants submit to the market, no significant difference was found between orders submitted through stationary or mobile devices (Hypothesis 1b). As described in Section 7.3.2, both measures (*quantity* and *market order*) can be interpreted as proxies for confidence. Hence, it has to be concluded that the decision confidence seems to be unaffected by the device used.

TABLE 7.1: *Regression Model for Hypotheses 1a and 1b*

Hypothesis	**H1a** Quantity	**H1b** Market Order
Device (mobile)	207.50	−.08
	(1.23)	(−.68)
Market Order	−50.20**	
	(−2.96)	
Trade direction	68.59***	.46***
	(3.98)	(39.96)
Quantity		−0.55**
		−2.94
Control for Products	✓	✓
(Intercept)	895.59***	.02
	(28.97)	(.85)
Adj. R^2	2.60 %	
pseudo-R^2		8.65 %
N	131,561	131,561

Notes: t-statistic (left model) and z-statistics (right model) in parenthesis; pseudo-R^2: Nagelkerke (1991); $^{\cdot}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

7.4.2 Trading Behavior and Performance

Following the research model, the effect of *Trading Behavior* on *Trading/Decision Outcome* is analyzed (see H2 in Figure 7.1). Therefore, it is regard how *profit* and *quantity* of orders are connected. Using Equation 7.5, the model on the right side of Table 7.2 is received.

TABLE 7.2: *Regression Model for Hypotheses 2b/3a and 2a/3b*

Hypotheses	**H2b/H3a** Score	**H2a/H3b** Profit
Device (mobile)	$-.29^{*}$	-263.64^{*}
	(-2.53)	(-1.97)
Market Order	-0.63^{***}	121.73^{***}
	(-5.42)	(8.98)
Trade direction	$.74^{***}$	1655.26^{***}
	(62.48)	(120.56)
Quantity	$-.00^{***}$	$-.02^{***}$
	(-17.81)	(-8.52)
Control for Products	✓	✓
(Intercept)	$-.13^{***}$	-766.70^{***}
	(-6.02)	(-31.00)
Adj. R^2		10.68 %
pseudo-R^2	11.08 %	
N	131,561	131,561

Notes: t-statistic (left model) and z-statistics (right model) in parenthesis; pseudo-R^2: Nagelkerke (1991); $^{\cdot}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

As one can see, *quantity* has a negative significant influence on *profit*. However, since the effect strength is relatively small in comparison to the average profit (mean: 198.50, sd: 2,561.52), it is doubtful that there is an economic significant effect of *quantity* on *profit* and Hypothesis 2a cannot be proven. Turning to Hypothesis 2b, the model on the left side of Table 7.2 is the result of applying Equation 7.4 to the dataset. *Market order* has a small (marginal effect: −.015) negative influence on *score* and therefore supports Hypothesis 2b. Interestingly, although *market order* has a small negative influence on *score*, it has a positive influence on the overall *profit*, which leads to the conclusion that market orders fail at being profitable more often than their counterparts, but if they do not, they are much more profitable than limit orders.

Finally, the difference in Decision Outcome between the two interface types is regarded (see H3 in Figure 7.1). As described, two proxies are used to measure decision outcome in the market environment. First, the likelihood of an order being profitable (*score*) in regards to the *device* the order was submitted from (see Hypothesis 3a as depicted in Table 7.2). As one can see, *device* negatively affects this likelihood, although the regression controlled for *Trading Behavior* (*quantity*, *market order*) and market specific effects (Equation 7.3). Although the influence of *device* is rather small (marginal effect: $-.07$) and only significant at the 5 %-level, this result supports Hypothesis 3a. Moreover, *device* significantly reduces the average *profit* by about 264 currency units (see Hypothesis 3b in Table 7.2), and hence supports Hypothesis 3b.

As both the likelihood for submitting a profitable order (i. e., *score*) as well as the average *profit* decreases if the *device* dummy is set to 1, it can be concluded that submitting an order through a mobile device leads to worse trading performance compared to a stationary device. However, as this is a field experiment it cannot be identified to which extent this effect is driven by the mobile device, or by the environment in which a participant is trading with the mobile app (or even the mobile app itself). So both the mobile device and the environment in which participants use their mobile device may negatively affect their performance. However, it is not possible to give an explicit answer with the information available in this field study.

7.5 Conclusion

By describing a prediction market which participants can access through stationary and mobile interfaces the potential of analyzing decision processes in various device settings has been shown. Participants' confidence and trading behavior on a per-order basis has been examined subject to the device used. Furthermore, the influence of trading behavior, and the usage of a specific device class (namely, web or mobile) on trading outcome has been analyzed. This study contributes three main findings:

First, it was not possible to proof that the device has an influence on participants' decision confidence as measured by two proxies (order-size and order-type). Since decision confidence is only one aspect of what can possibly be influenced by using a certain device type, this study leaves room for further research.

Second, it could be shown that market orders, although they tend to lead to a higher profit, lead to profitable transactions less often. These somewhat contradictory results prompt the following interpretation: market orders fail more often, but if they do not, the average gain is higher than the average loss is in case of failing. In order to deeply understand trading behavior in this market, one has to analyze the trading behavior even more thoroughly, e. g., by taking additional aspects of trading behavior into account or simply by analyzing the data on a per-user basis.

Third, it has been found that orders submitted by a mobile device perform significantly worse than their stationary counterparts. In particular, significantly lower profits, and a significantly lower probability for submitting a profitable order were found when using a mobile device. It remains unclear, if the usage of the device itself or indirect influences are causing this 'performance penalty'. One might assume that mobile traders are simply distracted through an often-noisy environment. Another possibility is that mobile traders are simply unable to obtain information necessary to perform well via the smaller screen and other limitations of the user interface. As hinted at earlier, it is possible that the information usage as well as the environmental influence is worth considering in a future study.

A major limitation of this study is, that it can neither be ruled out that – in some cases – traders used the web-interface in a mobile setting to submit an order (e. g., with a laptop or a smartphone web-browser) nor that a user in a stationary setting used the mobile application for that purpose. Even though there are reason to believe that traders in a stationary context prefer to use the web-interface as well as mobile traders tend to use the KAPP, there are no reliable information about a user's trading environment. Hence, further research and a different, more controlled approach is needed to provide clarification of these questions. Moreover, the data has a strong bias towards orders submitted via the web-interface. One explanation for the few mobile orders could be that many participants tested the mobile interface with a couple of orders once and decided to use the stationary interface instead. Unfortunately, no acceptance-survey was conducted amongst users of KAPP. As the majority of feedback received concerning the usability of KAPP was positive, there is reason to believe that the few mobile orders are instead attributed to the participants' usage preferences of that particular market. In other words: Participants simply made little use of the possibilities of mobile trading. Besides this weakness, all orders submitted via mobile devices are real world observations traders submitted without being specifically incentivized to do so.

This study has two major implications: First, it illustrates that decision making performance does not solely depend on the decision maker and her resources. Second, one need to be aware of these differences when designing software artifacts using multiple devices having different characteristics. Specifically in the domain of financial markets this study is the first work to highlight the influence of mobile trading interfaces on trading behavior and performance. Due to the close relation to decision processes, this study helps to understand the impact of information system interfaces on decision making in general.

Chapter 8

Improving the *(IT) Infrastructure*: Interface Influence on the Disposition Effect

> " Traders in any, perhaps all, markets have different talents, interests, and abilities; they may interpret data differently or be swayed by fads. However, as long as not all traders are so influenced there is room for markets to function efficiently."
>
> ROBERT FORSYTHE, FORREST NELSON, GEORGE R. NEUMANN, AND JACK WRIGHT, 1992

8.1 Introduction

IN 2002, 28 % of U.S. retail trades were executed via retail brokerage companies; one year later, U.S. online retail brokerage companies already managed more than 31 million accounts (Bakos et al., 2005). In a more recent analysis Camargo and Fonseca (2013) estimate the US self-directed online brokerage market to have reached over 40 million customers in 2012. Furthermore, they report growth rates have slowed down since 2010 which can be an indication of saturation. From a customer's perspective, important distinguishing features for online retail brokerage companies are fees, trading capabilities, and the functionality of their trading interfaces. As designing trading interfaces gives brokerage companies an additional opportunity to differentiate themselves from competitors, a lot of effort is put into designing 'attractive' trading interfaces for customers. Although it is common knowledge, that a decision maker's performance in general depends inter alia on the user interface used (Speier and Morris, 2003), even a carefully designed and easy to use interface may not prevent traders suffering from behavioral biases.

The *disposition effect* is such a behavioral bias, which often leads to individual losses and missed gains. Although the disposition effect is well known in several research communities, it is not considered to be part of general knowledge. Therefore, providers of online trading platforms might have a particular interest to inform their customers about that bias, -and if possible provide tools to avoid the bias. On the one hand, the strength of that bias is influenced by the individuals' internal decision making processes and knowledge about the specific bias and awareness. On the other hand, the individuals' environment (e. g., information presentation) might impact the effect strength.

In this study, performance indicators are identified as a driver of the disposition effect and it is shown that their disuse can decrease the disposition effect and therefore its negative implications. Although, the question persists how individuals can be sensitized for this bias. Evidence that textual information can work – even under difficult circumstances – can for instance be found in the area of health warnings on tobacco. Hammond (2011) could show that persons who noticed a textual warning sign, in some cases started to think about changing their behavior. But he also emphasizes, that the information must *"capture [. . .] attention and educate"* (Hammond, 2011) in order to be effective. Another study in the health domain examining effects of pictures and textual information found that only using textual arguments led to minor changes in intended behavior (Boer et al., 2006). To summarize, it has been shown that textual information can have an effect, although it does not seem to be a strong one. The following research questions are addressed in this study: (i) Is the knowledge about the existence of the disposition effect suitable to lower the disposition effect exhibited by an individual? (ii) Does a trend indicator arrow (like the ones often used in online trading screens) positively affect the strength of the disposition effect exhibited by an individual? To answer these questions, an experiment is set up in an online prediction market. Before addressing these questions it is first verified that the disposition effect is prevalent in the regarded market at all, and whether it effects participant's trading performance.

The remainder of this chapter is structured as follows: First, related research concerning the disposition effect and prediction markets is presented in Section 8.2. Second, the conducted experiment is described and the hypotheses are developed in Section 8.3. Thereafter, in Section 8.4, a short description of the dataset is given and the methodology used is outlined, before the findings are presented. Finally, results and their implications are discussed and concluding remarks are made in Section 8.5.

8.2 Related Work

8.2.1 Disposition Effect

Across a wide range of markets, traders tend to hold on to paper losses for too long and realize gains too early. This tendency is a deviation from rational behavior, where the trader makes his decision based on relative gains and losses instead of the absolute valuation of his investment. Based on Kahneman and Tversky's (1979) prospect theory, the work of Machina (1982), and others, Shefrin and Statman (1985) examined this particular pattern and coined the term *disposition effect* (DE) for it. They developed a descriptive theory that enabled a broader insight on this particular effect in real markets. But their explanatory approach goes beyond prospect theory and also includes aspects of mental accounting (Thaler, 1985), as well as the asymmetry of pride and regret (Kahneman and Tversky, 1979; Thaler, 1985), and self-control (Thaler and Shefrin, 1981). The existence of the disposition effect has been shown in stock markets (e. g., Lakonishok and Smidt, 1986), for a U.S. discount brokerage house (Odean, 1998) or for the Taiwan Stock Exchange (Barber et al., 2007), but also in experimental settings (e. g., Andreassen, 1988; Weber and Camerer, 1998) or in prediction markets (e. g., Teschner et al., 2012). Although the disposition effect can be shown in a wide range of markets, its strength seems to depend on individual factors, such as professionalism, sophistication, and trading experience. Shapira and Venezia (2001) examine a dataset from an Israeli brokerage house and found out, that independent investors tend to have a higher disposition effect than professional investors. Seru et al. (2010) show that the disposition effect declines with trading experience. But even a lower disposition effect for professional traders does not mean, that disposition effect's performance-degrading implication vanishes with growing experience. Both studies imply that the strength of the disposition effect for an individual is varying and can actively be influenced. Garvey and Murphy (2004) analyzes a successful team of proprietary traders and found, that even though the traders were experienced and performed very well, their performance could have been better, if they would have avoided the disposition effect's trading pattern. Feng and Seasholes (2005) show, that a combination of sophistication and trading experience can even eliminate investor's reluctance to realize losses, but it can only diminish the propensity of an investor to realize gains. Summing up, the disposition effect has shown to hinder individuals trading performance. Although, it can be diminished by traders' experience and sophistication it cannot be totally avoided.

8.2.2 Disposition Effect in Prediction Markets

Teschner et al. (2012) analyzed the disposition effect in a prediction market for macroeconomic indicators as described in Teschner et al. (2011) with a sample size of 96 active traders. They conducted their analysis largely based on the work of Odean (1998). In line with previous research, they found a disposition effect on the individual level (DE = .1582) as well as on the aggregated level (DE = .2248). Furthermore, they found a significant asymmetry in the disposition effect towards the percentage of gains realized. Interestingly, there was no significant impact of the disposition effect on absolute forecast error as well as no correlation between prediction accuracy and disposition effect. Hartzmark and Solomon (2012) examined a dataset of a NFL betting market from Tradesports.com, Inc.[1] and found that prices followed a S-shaped curve instead of linearly matching the underlying probabilities. They found this particular mispricing to be consistent with the disposition effect. In another study, Borghesi (2013) found strong evidence for the disposition effect in Tradesports' market for NBA totals contracts to lead to significant differences between prices and underlying values, also consistent with the disposition effect. Summing up, there is evidence that the disposition effect exists in prediction markets.

8.3 Setting and Research Questions

8.3.1 Experimental Setting

A field experiment is conducted on a prediction market called Kurspiloten (cf. Section 4.2). Additional specifics of that market are detailed in the next paragraph. Afterwards, the ex-

[1]Accessible at the URL `http://www.tradesports.com/` .

periment, consisting of treatment-specific user interface changes is described. The treatments actually used are explained in the last paragraph.

Market Details

Besides individual decisions, the disposition effect also depends on market price developments. For example, traders in a bearish market have simply less chances to realize paper gains but more paper losses; the opposite applies to bullish markets. Since traders might concentrate their trading on different stocks, this dependency might be problematic for further analyses. In extreme situations traders might experience different or even opposed market effects due to their different portfolios. However, it is expected that the effect market price developments have on the disposition effect are rather small in this market. First, the tradable stocks (Table 4.1) can roughly be grouped into stock indices and commodities. Within those groups, the single commodities/indices are somehow interdependent (e. g., DAX and MDAX, Gold and Silver.) and thus are unlikely to develop in opposed directions for a longer period of time. Second, traders in Kurspiloten market start with an identical portfolio and receive an identical endowment each week, therefore tempting traders to trade all kinds of tradable stocks. As all traders participate in the very same market, it is assumed that price market trends do not influence the disposition effect between individual traders significantly. Finally, traders' profits are used as a control variable in the following regression models (where appropriate), which further smoothens the potential negative impact of market price developments on the comparability of the individual disposition effect.

User Interface Modifications

The experiment is set up as a 2 × 2 full factorial between subjects design. Both treatment conditions are visual changes to the trade screen (Figure 8.1; Appendix A). The first change ('DE Info Text') consists of a linked text "Do you know about the disposition effect?"[2] just above the price chart (see label (a) in Figure 8.1). When a user clicks on this text, a paragraph explaining the disposition effect fades in. Appendix A contains the complete text besides an english translation. As the experiment takes place in the field, compromises must be made in some areas. Hence, traders are not forced to read the DE Info Text prior to trading on the market. Instead the current time and user id is recorded with every click

[2] Author's translation. Original phrase: "Kennen Sie den Dispositions-Effekt?"

on the link to DE Info Text for further analyses. The second treatment condition ('Trend Indicator') extended the box "Your Performance"[3] on the lower right of the trade screen by one column (see label (b) in Figure 8.1). The basic interface only contains the information "Average Purchase Price"[4] (left column in the box "Your Performance"), whilst the second treatment condition extended that box by a column named "Performance", containing the relative performance of stocks held. First, the percentage difference between the current market price and the average purchase price for the corresponding stock is shown. Second, a tiny trend direction arrow indicates whether this difference is negative, zero, or positive. The arrow is colored red, grey, or green, respectively. It is similar to stock trend indicators used in many online trading interfaces.

Treatments

All participants registered on the Kurspiloten market are assigned to one of the three treatment groups or to the control group as shown in Table 8.1. Participants who registered in the pre-market phase are randomly assigned prior to the start of the market. Participants who joined after start of the market are assigned randomly at registration. Each trader remains member of the assigned treatment group for the whole duration of the market. The first group was confronted with both conditions described above (treatment *Trend_Info*) and depicted in Figure 8.1. One group saw the trend info (treatment *Trend*), another one the info text (treatment *Info*). No changes were made for the control group (*Control*), i. e., the control group saw neither the info text nor the trend info.

TABLE 8.1: *Treatments and Research Design*

	DE Info Text	**w/o DE Info Text**
Trend Indicator	*Trend_Info*	*Trend*
w/o Trend Indicator	*Info*	*Control*

8.3.2 Research Questions

This study tries to answer the afore-mentioned research questions 5 and 6 from Section 1.2:

[3] Author's translation. Original phrase: "IHRE PERFORMANCE"
[4] Author's translation. Original phrase: "durchschnittlicher Kaufpreis"

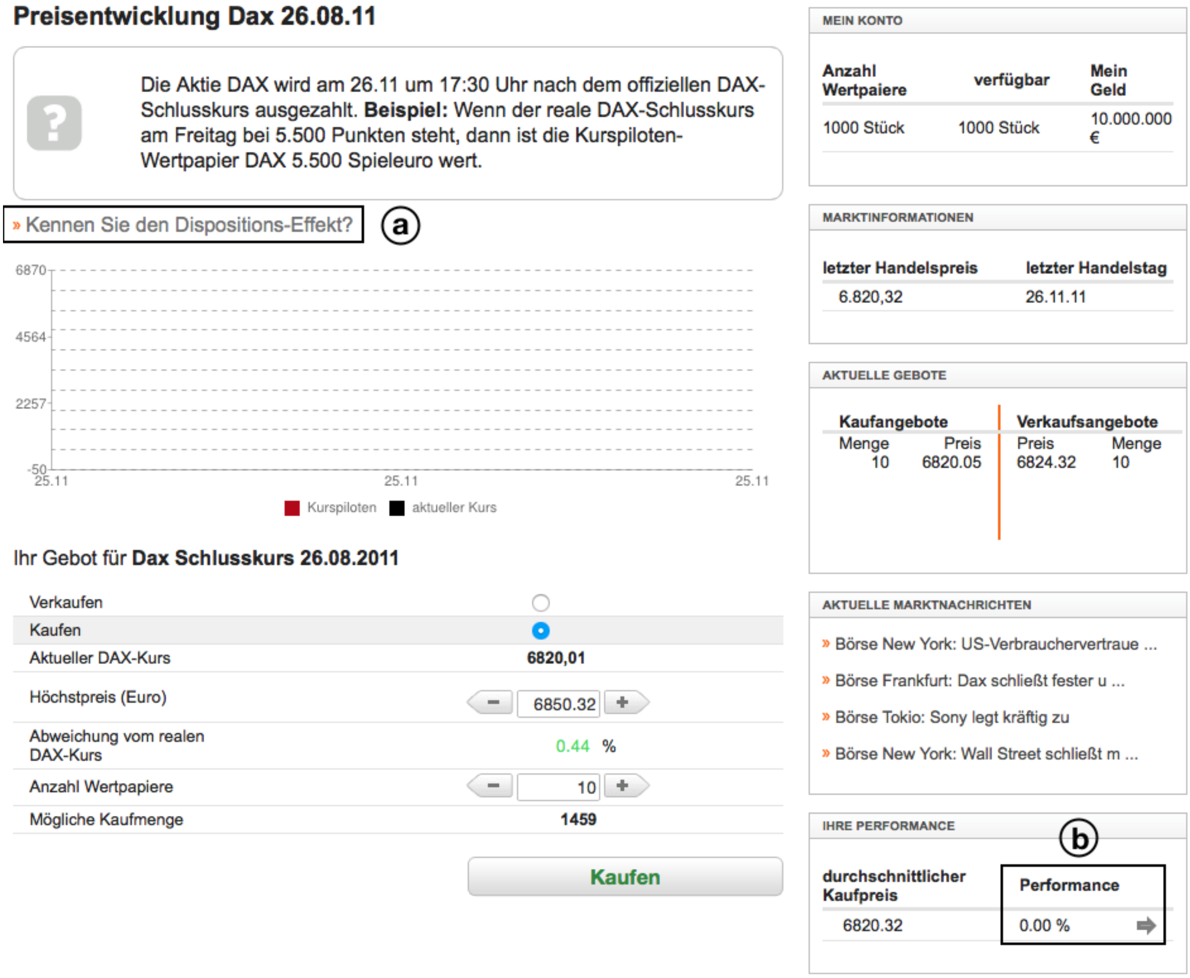


FIGURE 8.1: *Trading Screen for Treatment Trend_Info*

(Containing both user interface modifications made: a and b. A click on the linked text (a) fades in an info text about the disposition effect. The whole disposition info text is depicted in Appendix A. Modification (b) shows the 'Trend Indicator' element. Screenshots of the three remaining treatments can be found in Appendix A.)

Heading: "Price development of Dax 2011-10-07"; In box (a): "Do you know about the disposition effect" (only available in treatments *Info* and *Trend_Info*); Chart: price chart for *Kurspiloten* prices (red dotted line) and real-world prices (black line); middle left: "Your Order for Stock ... 2011", radio buttons for buy and sell, information about the current real-world price of selected stock (bold), input field for limit price, information about deviation of limit price from real-world price, input field for quantity, information about buying power (bold), 'execute' button; Right column: 1^{st} box: "My Portfolio", own holdings, own holdings available, and money (P€); 2^{nd} box: "Market Information", least recent price and closing date of current product; 3^{rd} box: "Orderbook"; 4^{th} box: "Current News", news stream from a major German financial newspaper; 5^{th} box: "Your Performance", average purchase price of selected stock and relative performance (i. e., relative price difference of average purchase price and least recent market price; only available in treatments *Trend* and *Trend_Info*)

Research Question 5: *Is providing information about the disposition effect suitable to lower the strength of the disposition effect exhibited by an individual?*

Research Question 6: *Does a trend indicator arrow affect the strength of the disposition effect exhibited by an individual?*

As stated earlier, albeit the disposition effect (DE) is a well-known behavioral bias, it is not part of a general education and can therefore not expected to be known by the vast majority of participants of an online prediction market. As research has shown, knowledge about the existence of the disposition effect can lead to a decreased disposition effect. This study tries to shed some light on the question, if it is expedient to inform about the DE with a short information text directly within an online trade screen or if a 'deeper understanding' of the disposition effect is needed. It is expected that reading an information text leads to a lower disposition effect, simply by creating awareness for this particular deviation from rationality, and thus increasing self-control.

Therefore, an – yet unspecified – interface change 'DE Info Text' is defined, consisting of an information text about the disposition effect on the trading screen. Hence, in line with current research, the information text is expected to reduce the disposition effect:

Hypothesis 4: *Mean disposition effect is lower if 'DE Info Text' was read. (INFO < CTRL)*

Moreover, self-control might be decreased by confronting a trader with the portfolio state in a transparent fashion.The disposition effect is driven by the traders' perception of his portfolio development; e. g., if a trader cannot remember the purchase price of stocks, he is obviously unable to tell if he is riding a gain or a loss. In a more complex market environment, traders repeatedly buy and sell different amounts of shares for different prices resulting in a non-intuitive way to calculate the average purchase price. That purchase price has to be compared to the current stock market price in order to determine the own holdings' performance. The easier a trader realizes his portfolio value, the more he might be tempted to yield to the disposition effect.

Furthermore, it is well known that traders can fall victim to mental accounting. Showing traders a transparent state of their portfolio on a per stock basis might intensify this biased perception. In order to support traders with an objective and comparable method to reflect about the portfolio performance, an interface change 'Trend Indicator' is defined, that consists of a relative performance indicator of a trader's portfolio price development and a visual cue representing its direction. The 'Trend Indicator' is therefore expected to increase the disposition effect:

Hypothesis 5: *Mean disposition effect is higher if 'Trend Indicator' is present. (TREND > CTRL)*

From a theoretical point of view, it is not expected that an information text about the disposition effect and a trend indicator for the own portfolio's price development should influence each other. Therefore, the following hypotheses are developed:

Hypothesis 6a: *'Trend Indicator' does increase the mean disposition effect, even if 'DE Info Text' is present. (TREND_INFO > INFO)*

Hypothesis 6b: *'DE Info Text' is suitable to reduce the mean disposition effect, even if 'Trend Indicator' is present. (TREND_INFO < TREND)*

The trend indicator is expected to induce a higher order activity, since it reflects the state of a portfolio in a more transparent way and thus might make trading opportunities more obvious. Hence, the fourth hypothesis is:

Hypothesis 7: *'Trend Indicator' leads to an increase in the traders' activity.*

8.4 Results

In this section the empirical findings are presented, starting with descriptive statistics. Afterwards, the overall and individual existence of the disposition effect is shown, before a detailed look at the disposition effect with regard to the four treatments introduced earlier is taken. Finally, we shed some light on the traders' order-based activity per treatment.

8.4.1 Descriptive Statistics and Methodology

This study uses the dataset from Kurspiloten market as described in Section 4.2. The disposition effect is only calculated for traders who submitted at least 12 orders. Additionally, traders that had no chance to realize a gain or a loss and traders that did not realize at least one gain or one lose are filtered. Due to these circumstances the disposition effect can be determined for 514 traders. The sizes of the three treatment groups and the control group are nearly balanced out: $N_{\text{Trend\_Info}} = 123$, $N_{\text{Trend}} = 126$, $N_{\text{Info}} = 123$, and $N_{\text{Control}} = 142$. About one quarter of the traders, who could click on the info text link, actually made use of this possibility: $N^{\text{clicked}}_{\text{Trend\_Info}} = 30$, $N^{\text{clicked}}_{\text{Info}} = 30$. The average account age lies between 69.10 days (*Trend*) and 67.26 days (*Trend_Info*) with an overall mean of 68.05. Traders' performance – measured by their total trading profit – differs significantly (t-stat = 2.34, p = 2.04 %) between treatments *Info* and *Trend*. Hence, variable *Profit* is used in the regression analysis to control for that fact. Besides, variable *Trades per Day* is used to control for different trading activity. Although the number of trades per day does not significantly differ between treatments, it does for traders that clicked on the info link in comparison to those who did not (t-stat = 2.51, p = 1.24 %).

The disposition effect is mainly measured based on Odean (1998). The only exception is the length of the time slices used. Since the trading period per product was rather short (seven days), the users' sessions are used to differentiate between paper gains and losses instead of trading days; e. g., if a traders' average purchase price was below the highest and lowest market price in the regarded session it is counted as a paper gain. The disposition effect (DE) is calculated as DE = PGR−PLR where PLR denotes the Proportion of Losses Realized, and PGR the Proportion of Gains Realized. PGR and PLR are calculated as follows:

$$\text{PLR} = \frac{\#\ \text{realized losses}}{\#\ \text{realized losses} + \#\ \text{paper losses}} \tag{8.1}$$

$$\text{PGR} = \frac{\#\ \text{realized gains}}{\#\ \text{realized gains} + \#\ \text{paper gains}} \tag{8.2}$$

8.4.2 Disposition Effect on Prediction Markets

In line with current research, an aggregated disposition effect (DE) can be shown in the Kurspiloten market (DE = .154, PLR = .041, PGR = .196) which is slightly smaller than in a similar study of Teschner et al. (2012) (DE = .225, PLR = .018, PGR = .242) and higher as in studies using data of online brokers (e. g., for an U.S. discount broker Odean (1998) measured DE = .05, for a German online broker Weber and Welfens (2007) measured DE = .09). On the individual level the disposition effect is 0.148 and thus comparable to a similar study on a play-money prediction market conducted by Teschner et al. (2012) (DE = .158). Further details are displayed in Table 8.2. As one can see PLR, PGR and DE are significantly greater than zero. Additionally, the disposition effect is asymmetric, since the absolute correlation between DE and PLR is slightly smaller than between DE and PGR.

TABLE 8.2: *Mean Individual Disposition Effect*

	Value	**t-stat** ($x > 0$)	**Correlation** (DE, x)
PLR	.094	11.67	−.69
PGR	.242	21.84	.85
DE	.148	9.89	—

Notes: N = 514 (complete groups); both correlations are significant at a 1 %-level

Result 3: *The disposition effect is prevalent in the regarded market on an aggregated as well as on an individual level.*

8.4.3 Disposition Effect's Influence on Trading Performance

As the disposition effect is prevalent in the market, the question arises, how the disposition effect influences the market. Since this study focuses on the trader, there is a particular interest in the disposition effect's influence on traders' performance. Therefore, the correlations between the traders' profits and the disposition effect, as well as their relative rank and the disposition effect are regarded. (Relative rank here indicates the rank within the 514 regarded traders instead of the overall rank among all registered traders.) Neither a significant correlation between profits and the disposition effect ($\rho = .011$, Pearson's product-moment correlation, t-value = .24), nor between the disposition effect and the traders' rank ($\rho = -.028$, Pearson's product-moment correlation, t-value = −.62) can be found.

8.4.4 Disposition Effect per Treatment

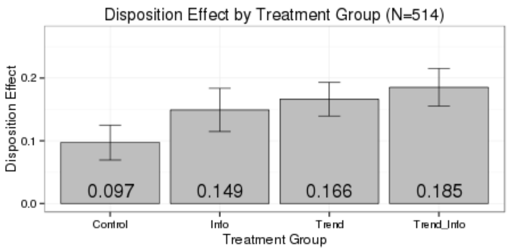
Table 8.3 (Figure 8.2a) shows the mean disposition effect in each treatment group and the control group. The differences between *Trend_Info* and *Trend* ($\delta = .019$), *Info* and *Control* ($\delta = .052$), *Trend_Info* and *Info* ($\delta = .036$), and Trend and Info ($\delta = .017$) are not significant. Solely, the disposition effect for *Trend_Info* as well as for *Trend* is significantly higher than for *Control* (both on a 5 %-level; *Trend_Info:* $\delta = .088$, t-stat = 2.15, p-value = .016 and *Trend*: $\delta = .069$, t-stat = 1.78, p-value = .038). At first glance, this result seems to support Hypothesis 5. But as mentioned earlier, even if all traders in treatments *Info* and

Trend_Info may read the disposition effect info text, it has not been controlled whether they actually *did* expand this info text yet.

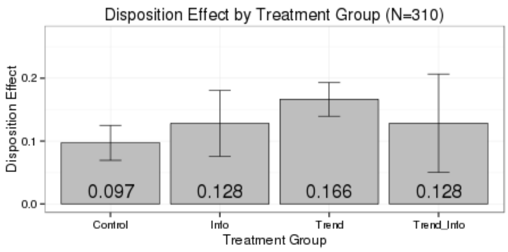
TABLE 8.3: *Mean Individual Disposition Effect per Treatment (Complete Groups)*

	DE Info Text	**w/o DE Info Text**
Trend Indicator	*.185*	*.166*
w/o Trend Indicator	.149	.097

Notes: N = 514 (complete groups)



(A) *complete groups*



(B) *subsample*

FIGURE 8.2: *Mean Individual Disposition Effect per Treatment*

Therefore, the former analysis is repeated with a slight adaption: Table 8.4 (and Figure 8.2b) shows the mean disposition effect for a subsample, in which only traders in the *Info* and *Trend_Info* treatment were taken into account that clicked on the info text link. Furthermore, the disposition effect for those traders is calculated on the basis of trades they executed *after* they first clicked on the info text link. Hence, N is slightly smaller.

As one can see, there is hardly difference between the treatments *Trend_Info* and *Info* (δ = .001, not sign.). Also, the differences between *Trend_Info* and *Trend* (δ = .038), *Info* and *Trend* (δ = .038), *Trend_Info* and *Control* (δ = .031), as well as *Info* and *Control* (δ = .031) are not significant. Solely, the disposition effect for *Trend* is significantly higher than for *Control* on a 5 %-level (δ = .069, t-stat = 1.78, p-value = .038). Again, this finding supports Hypothesis 5. However, theses results cannot confirm Hypothesis 4.

Result 4: *Textual information about the disposition effect has no influence on its strength (cf. Hypothesis 4).*

Result 5: *Treatment Trend shows a significantly higher disposition effect than the control group (cf. Hypothesis 5).*

TABLE 8.4: *Mean Individual Disposition Effect per Treatment (Subsample)*

	DE Info Text	w/o DE Info Text
Trend Indicator	.128	*.166*
w/o Trend Indicator	.128	.097

Notes: N = 328 (subsample: all traders who have not clicked on the 'DE Info Text' link mentioned in subsection 'User Interface Modifications' are filtered.)

When taking a look at the tiny difference between *Trend_Info* and *Trend* ($\delta = .001$) in Table 8.4 in contrast to the rather big difference of 0.069 between *Trend* and *Control*, one might assume that the treatment condition 'DE Info Text' might have an influence on the treatment condition 'Trend Indicator'. It seems reasonable to examine, if the info text does hinder the trend indicator's increasing influence on the disposition effect. To control for such an interaction effect, the regression model shown in Table 8.5, Model 2 and 4 is used. To control for such an interaction effect, the regression model shown in Table 8.5, Model 2 and 4 is used.[5] Additionally an ANOVA was applied. Neither method shows an interaction between the treatment conditions 'DE Info Text' and 'Trend Indicator'. That means, neither of the treatment conditions do have a stronger or weaker effect under the premise that the other treatment condition is present or absent. Furthermore, Models 3 and 4 control for potential differences in treatment groups. Nevertheless, all models in Table 8.5 show a positive influence of the trend indicator on the individual disposition effect.

For the sake of completeness, Table 8.6 contains the result for the complete treatment groups. The result of the OLS regression and the ANOVA are qualitatively similar to what has been presented for the subsample (Table 8.5), including the trend indicator's influence on the disposition effect.

Result 6: *No interaction effects have been found between the disposition effect and showing visual cues (cf. Hypotheses 6a).*

There is reason to believe, that the trend indicator itself does increase the individual disposition effect strength. Therefore, the average disposition effect of all traders who can see the trend indicator (mean DE = .176) is compared, with those who cannot (mean DE = .121). In other words, the joined treatment *Trend* and *Trend_Info* is compared against treatment *Info* and the *Control* group. This analysis results in a significantly higher

[5] Please note, that the dummy-coding was adjusted for 'DE Info Text' appropriately.

TABLE 8.5: *Measuring Interaction Effects (Subsample)*

Model	**(1)** Direct Effects	**(2)** Interaction	**(3)** (1) + Controls	**(4)** (2) + Controls
Trend Indicator	$.060^{\cdot}$	$.069^{\cdot}$	$.061^{\cdot}$	$.070^{\cdot}$
(visible=1, hidden=0)	(1.66)	(1.78)	(1.70)	(1.81)
DE Info Text	−.05	.031	−.014	.020
(visible=1, hidden=0)	(−0.09)	(.41)	(−.27)	(.27)
Trend Indicator		−.069		−.067
×		(−.65)		(−.63)
DE Info Text				
Trades			$.003^{\cdot}$	$.003^{\cdot}$
per Day			(1.89)	(1.87)
Profit			.000	.000
			(−.35)	(−.36)
(Intercept)	$.102^{***}$	$.097^{***}$	$.088^{***}$	$.084^{**}$
	(3.96)	(3.66)	(3.33)	(3.07)
Adj. R^2	.25 %	.06 %	.76 %	.56 %
N	310	310	310	310

Notes: OLS regression estimates on subsample; dependent variable: disposition effect; t-statistics in parenthesis; $^{\cdot}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

value for traders who see the trend indicator on a 5 %-level (δ = .054, t-stat = 1.82, p-value = .034, N = 514). Again, the repetition of this analysis for the subsample from above leads to analogous results (δ = .060, t-stat = 1.66, p-value = .049, N = 310).

> **Result 7:** *Displaying visual cues such as trend arrows increases the individual disposition effect (cf. Hypothesis 5).*

8.4.5 Activity per Treatment

As mentioned in section Hypotheses, treatment *Trend* is expected to have a higher number of orders submitted (Hypothesis 7). Therefore the number of (i) orders submitted and (ii) logarithmized number of orders submitted is compared between all treatment groups. The logarithmization is used, since it diminishes the effect of extreme values.

TABLE 8.6: *Measuring Interaction Effects (Complete Groups)*

Model	**(1)** Direct Effects	**(2)** Interaction	**(3)** (1) + Controls	**(4)** (2) + Controls
Trend Indicator	.054˙	.062˙	.055˙	.062˙
(visible=1, hidden=0)	(1.83)	(1.94)	(1.83)	(1.95)
DE Info Tex	−.017	.014	−.025	.005
(visible=1, hidden=0)	(−.37)	(.21)	(−.54)	(.08)
Trend Indicator		−.062		−.061
×		(−.67)		(−.66)
DE Info Text				
Trades			.002	.002
per Day			(1.57)	(1.56)
Profit			.000	.000
			(−.29)	(−.31)
(Intercept)	.123***	.120***	.114***	.111***
	(5.75)	(5.43)	(5.16)	(4.87)
Adj. R^2	.28 %	.17 %	.37 %	.26 %
N	514	514	514	514

Notes: OLS regression estimates on subsample; dependent variable: disposition effect; t-statistics in parenthesis; $^{\cdot}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

One significant difference is found between *Info* and *Control* ($\delta = 382.24$, t-stat = 1.71, p-value = .045) for comparison 'i' and four differences for case 'ii': logarithmized trading activity in treatment *Info* is slightly higher than in *Control* group on a 0.1 %-level ($\delta = 1.126$, t-stat = 3.91, p-value < .001). Additionally, the logarithmized trading activities for treatment *Trend_Info* and *Info* are significantly higher than for treatment *Trend* (*Trend*: $\delta = 1.059$, t-stat = 3.55, p-value < .001, *Trend_Info*: $\delta = .819$, t-stat = 2.81, p-value = .003). Finally, the logarithmized trading activity for treatment *Trend_Info* is significantly higher than for *Control* ($\delta = .885$, t-stat = 3.15, p-value < .001).

Since this analysis uses the afore-stated subsample (only traders who clicked on the 'DE Info Text' link), the result may also be interpreted in a different way: the more orders a trader submits, the more often she sees the 'DE Info Text' link. Therefore one may argue, that it is more likely for her to click on this link, which will result in such a pattern. To

TABLE 8.7: *Activity per Treatment*

	DE Info Text	w/o DE Info Text
Trend Indicator	5.479	4.922
w/o Trend Indicator	5.502	4.855

Notes: N = 514 (complete groups); mean logarithmized number of orders submitted per Treatment

clarify that question, the complete group was analyzed, but no significant differences were found (Table 8.7). Hence, Hypothesis 7 is rejected.

Result 8: *The trend indicator does not lead to a higher trading activity (cf. Hypothesis 7).*

8.5 Conclusion

As the results show, the disposition effect can be aggravated by a tiny modification of the user interface. The modification consists of a simple percentage value and a trend direction arrow showing the traders' portfolio value, as used by online trading sites throughout the web as trend indicator for stock prices or for similar applications. Surprisingly, even such a small change does significantly increase the strength of the disposition effect. Those changes are not expected to only have a downside. On the upside, it is assumed that traders seeing the interface elements described above do submit more trades, since it shows the current state of the traders' portfolio in a fast and easy recognizable manner. But interestingly this assumption could not be verified. As private investors are regularly confronted to trading interfaces containing such elements, those results are especially interesting for providers of market interfaces. For market interface providers like retail brokers, these results imply to not use trend indicators. Nevertheless, currently most online brokers do make excessive use of such interface elements, at least for the reason of easier recognition of relative (price) changes. In order to help retail investors to avoid the disposition effect – which has previously been shown to reduce investor welfare – online brokers should consider redesigning their interfaces. These results also have an implication for regulators. They should carefully think about obligating online brokers to elucidate customers about behavioral biases, which are known to degrade their performance. As the results suggest, textual advices do not seem to be the best possible solution in this case. (Besides, the results put the effectiveness of textual information and advices already given to traders in question.) Regulators might furthermore think about banning certain types of visual cues that are known to lead to a great share of 'wrong' decisions and a substantial degradation of performance. Further research is needed to show, if the visual cue examined in this study does satisfy the requirements to belong into this category. Nevertheless, retail brokers should be interested in a good user experience and are hence motivated to deliver a 'good' user interface, which is supporting instead of misleading. Innovative retail brokers might even use results like these to create a unique sales proposition, playing a pioneer role in providing disposition effect-reducing user interfaces.

The disposition effect is a well-explored deviation from rational behavior. As many studies show, the disposition effect can have a negative impact on the decision performance in trading environments. This study analyzed the disposition effect on aggregated and individual level in an online prediction market with nearly 2,000 active traders and more than 200,000 orders. In line with research, a disposition effect could be found at both levels. Furthermore, a field study with over 500 traders was conducted for which the individual disposition effect could be measured. Although, it could not be verified that creating awareness of the disposition effect via textual information could decrease its strength, it could be shown that even tiny visual cues can significantly increase the strength of the disposition effect. Nevertheless, this study leaves room for further research. On the one hand, the trend indicator was solely used to represent the average purchase price of the traders' portfolios in contrast to the current market price. A future study could examine, if the disposition effect is also affected if trend indicators are used to represent price changes of tradable stocks. On the other hand, it has been reported, that only about one quarter of traders clicked on the link to the offered info text. Furthermore, even if a trader clicked on the link, there was no possibility to validate that she has (i) understood the text and its implications or (ii) read the text at all. A laboratory experiment could be set up to control for these factors; an additional questionnaire could provide certainty if a participant has read and understood the concept of the disposition effect and its implications on her trading performance. In a follow-up field study, a reposition of the offered link in a more conspicuous location is worth considering.

Chapter 9

Extending *Auxiliary Services*: Conducting Trader-centered Surveys

> "An understanding and appreciation of existing institutions, good theory, good computational modeling and well-designed experiments are critical ingredients to a successful design."
>
> HAL R. VARIAN, 2002

9.1 Introduction

ACCURATE and reliable forecasts of future short- and long-term events are a crucial competitive factor for companies, regions, and countries and an important foundation for political decision making. Advances in information systems are changing information aggregation in many contexts: political institutions increasingly open up for grassroots feedback and open discussion of societal innovation, ad-hoc communities use social media to coordinate, and companies gradually shift decisions towards a broad basis of employees and allow for user-driven innovation. An underlying theme of this trend is using the *collective intelligence* and *wisdom of the crowd*.

There are various ways to utilize the wisdom of crowds or collective intelligence such as using wikis, reputation systems, or polling mechanisms. Another way to aggregate dispersed information is by using a *Prediction Market* (cf. Chapter 3.) In these markets, participants trade contracts whose payoff depends on the outcome of uncertain future events. For example, a market contract might reward one dollar if a particular presidential candidate is elected. An individual who thinks the candidate has a 65 % chance of being elected should be willing to pay up to 65 cents for such a contract. Market participants form expectations about the outcome of an event. Comparable to financial markets, they buy if they find that prices underestimate the probability of the event in question and they sell a stock if prices overestimate the probability of an event. The track record of prediction markets suggests that markets may help to better foresee future developments and trends. Although, prediction markets have their strengths in quantitative predictions and even make conditional predictions possible – albeit complicated – (cf. Berg and Rietz, 2003), they are not well suited when it comes to qualitative predictions. The strength of prediction markets is the collaborative valuation of given contracts (i. e., the mapping between payoff, event outcome, and event date). Since all valuation is based on quantitative values (i. e., prices), qualitative information can only be induced into the market in the contract design, which falls to the market operator. For example, questions like “What will the Gross Domestic Product (GDP) of country A be in 2015” fit perfectly to prediction markets, whereas “How can productivity be improved?” is better suited for surveys.

Complex forecasts, such as *conditional* or *qualitative judgments* are better gathered with traditional forecast methods such as survey-systems. However, traditional survey-systems also have some known drawbacks. First of all, the success of surveys largely depends on the participant selection (Ammon, 2009; Gordon, 2007). The most common selection criterion is reputation, which is based on perceived expertise. However, Tetlock (2005) shows that perceived expertise does not correlate with individual forecast accuracy. The second drawback is the decreasing participant motivation over the study’s course. The long, rigid and tedious process leads to decreasing participant numbers (Cuhls, 2003). Ilieva et al. (2002) conducted a literature review and found response rates for online surveys from low as 6 % (Ranchhod and Zhou, 2001) to high as 67 % (Kiesler and Sproull, 1986). Deutskens et al. (2004) conducted a study with different types of surveys. Evans and Mathur (2005) analysed the pros and cons of online surveys in contrast to traditional mail surveys and discussed the online surveys’ best uses. Inter alia, he found online surveys are best to use if *timeline is vital, strong methodological control is sought* (e. g., order of the questions), and *survey research is conducted frequently*; all which applies for repeated online surveys in a prediction market context.

There are at least two ways surveys can benefit from an accompanying prediction market; *motivation* and *pre-selection of experts*. Prediction markets motivate participants to contribute continuously through incentives and by providing constant feedback; both on the aggregate and the individual level. More participants might be willing to participate (at least partly) in a survey if they have indicated that they have information regarding a topic. The question is how to figure out when a participant has information about a topic? This can be detected through the prediction market. If participants change the market price, they most likely have information about a certain topic and might be willing to fill out some related qualitative and possibly more complex questions. In previous studies (e. g., Chen et al., 2005), surveys run in parallel to prediction markets but in separate system. Survey participants had to leave the known platform, fill out a survey, and return to continue trading. Potentially, participants might find it convenient to answer survey questions right on the same platform. Moreover, Teschner et al. (2011) show that individual forecast input can be measured and objectively evaluated in prediction markets. Hence, this might help to pre-select experts *not based on their reputation* but on their previous *forecast performance*.

The remainder of this chapter is structured as follows: Section 9.2 gives a review of related work on prediction markets and surveys. The experimental setting and research questions are detailed in Section 9.3. Subsequently, Section 9.4 describes the dataset and the methodology used. Finally, Section 9.5 concludes this chapter.

9.2 Related Work

9.2.1 Prediction Markets and Surveys

Prediction markets offer a number of advantages over surveys. Prediction markets are continuous and ongoing, allowing immediate revelation of new information Rothschild (2009). As they are usually open around the clock, participants can trade whenever they like and therefore react to news immediately (Snowberg et al., 2007). This also applies mostly to (traditional) mail surveys (e. g., Dillman, 1991) or internet surveys (e. g., Zhang, 1999).[1] Although some surveys offer a small incentive in return for participation, the incentives earned by traders in a prediction market increase in proportion to the quality of

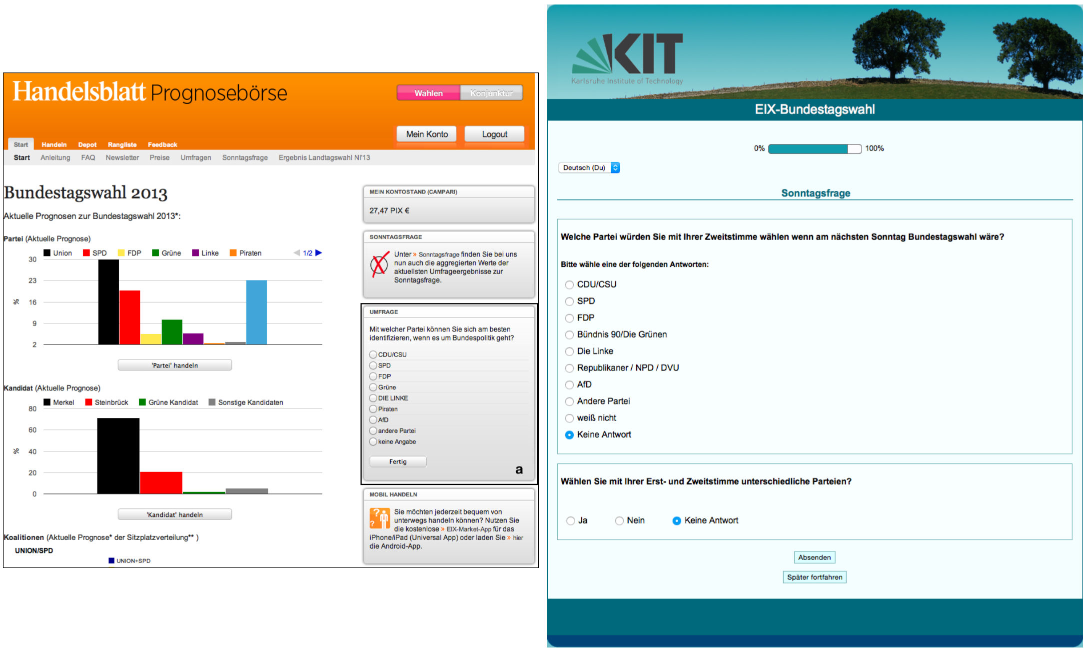
[1] See Cook et al. (2000) for an interesting overview on the response rates of internet surveys.

the information provided. Unlike surveys, a market provides immediate feedback to participants, allowing them opportunities to reassess their own information and to respond. The feedback enables participants to learn on two levels; first by actively trading, participants might gain experience and hence improve over time. Second, by observing their performance participants might realize their low ability and consequently leave the market (Teschner et al., 2011). The market interface is interactive and the setting gamified, in marked contrast to most surveys, providing further incentives for participation. Most surveys rely on random samples for validity and accuracy. In prediction markets, on the other hand, those with the best information are the best participants – the very individuals who are most likely to self-select into the market. Additionally, as successful participants accumulate their profits they gain forecasting weight over time compared to less successful participants. With surveys, this process of self-selection would introduce a sampling bias, but with markets, the incentive system forces low performers out of the market. Turning to the disadvantage of markets over surveys, one has to mention the higher complexity burdening participants (Graefe et al., 2010). First, they have to understand the trading mechanism and second, they have to understand how events are related to contracts. This process is more structured and better researched for surveys. The forecast performance of prediction markets is still in debate. On the positive side, they have proven repeatedly to be very potent information aggregation mechanisms (e. g., Berg et al., 2008; Ledyard et al., 2009; Bennouri et al., 2011). Although, other evidence suggests that the relative performance advantage of markets may be small compared to surveys or polls (e. g., Erikson and Wlezien, 2008; Rothschild, 2009; Goel et al., 2010). Prediction markets have a long track of successful field applications, e. g., in political elections (Berg et al., 2008), sport events (Luckner and Weinhardt, 2008), finance (Bennouri et al., 2011), and predicting market development (Spann and Skiera, 2003). See Wolfers and Zitzewitz (2004) and Ledyard et al. (2009) for reviews. However, to the best of the author's knowledge, prediction markets and surveys have never been combined during the prediction making process.

9.3 Setting and Research Questions

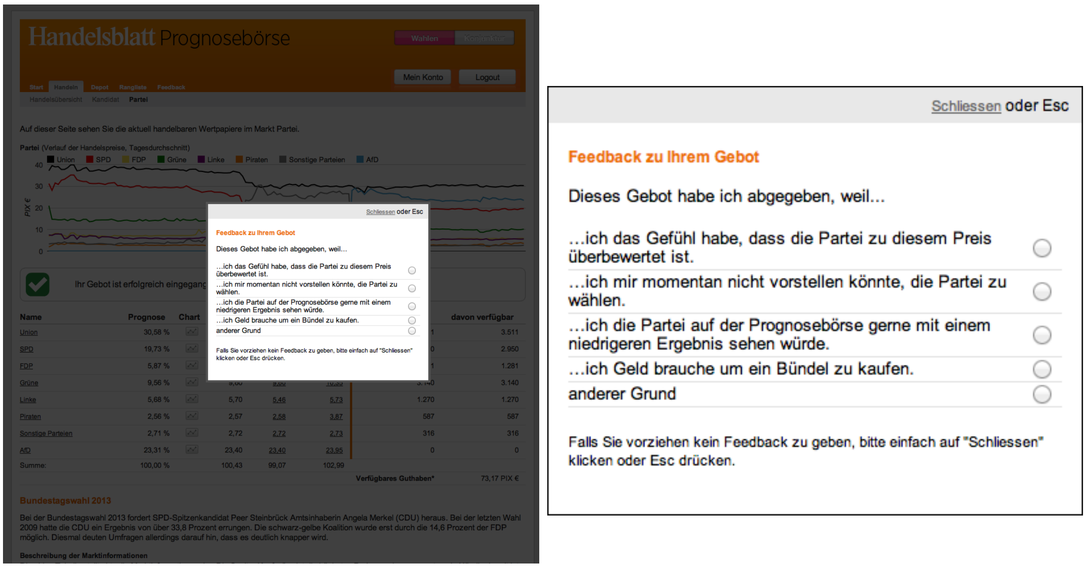
9.3.1 Experimental Setting

For this study a German political stock market, PIX, during the 2013 federal election hosted on the EIX market is used (cf. Section 4.3.1). In order to compare the integrated to standalone surveys two treatment groups (*integrated* vs. *standalone*) are set up. All traders are randomized in one of the groups. An extensive questionnaire with 73 items consisting of 6 parts was created: 1) General questions and platform feedback, 2) Election outcome, 3) Information sources, 4) Personal questions, 5) Political coalitions, and 6) Election Polls. While the integrated group could answer the questions one by one directly on the platform the standalone group was presented with a link to an external survey software (LimeSurvey, 2014) where participants had to answer all questions in one pass (Figure 9.1). The integrated questionnaire allows for questions with a predefined list of answers to appear on the PIX's main page, one at a time. Adjustable settings include the question mode, allowing only one reply selection or multiple. Equally important, it can precisely be defined *when*, *where* and *in what order* questions are asked. These properties are defined by a time window for question activity, and a priority ranking for the order of appearance. Obviously, it was made sure that every user can answer every question only once. Since these questions are of sensitive nature, each contains the option "prefer not to say". The answers can be selected via radio buttons, before users can submit their answers by clicking on a button. It was decided to use a self-made questionnaire infrastructure to enable traders replying quickly to a question without leaving the website. As a consequence, willing responders who simply not want to go to an external website were not lost.



(A) *Integrated Survey (box labelled 'a')* (B) *Standalone Survey*

FIGURE 9.1: *Integrated Survey (a) and Standalone Survey (b)*



(A) *Ad Hoc Post-Trade Question* (B) *Magnification*

FIGURE 9.2: *Trigger-based Survey on Top of Trade Overview Screen (a) and Magnification (b)*

Additionally, every participant is asked an 'ad hoc post-trade question' (*trigger-based survey*) using a self-made trigger-based survey-system. In order to gain a better understanding of the trader's thought processes, after every submitted order the trigger-based survey-system was called in order to decide, if the current participant should be asked the *trigger-based survey*. The algorithm used determines on the basis of the participant's trading history, number of already answered ad hoc post-trade questions in the last 21 days and a random component, if the question will show up immediately after the submission of an order or not. The parameters are configured in order to ask each trader at least once in three weeks per tradable product. The trigger-based survey asks the trader for the reason of his last order. Possible answers are: 1) *I feel, that party/candidate was under/overvalued*, 2) *I can/cannot imagine to vote for that party/support that candidate*, 3) *I like to see a lower/higher price for that party/candidate on this prediction market*, 4) *I need stocks/money to sell/buy a bundle*, 5) *other reason*.[2] The trigger-based survey (Figure 9.2) motivates participants to *rationalize* their trading decision ex-post and opens an interesting area for further field-studies on decision behavior.

[2]Note, that all answer were selectable via radio-buttons; especially, it was not possible to answer '5)' textual.

9.3.2 Research Questions

The integration opens up several research questions. First, how well (measured by response rate, total number of answered questions, response speed) does an integrated survey work compared to a standalone version? Surveys often have a quite low response rate of about 10 % (e. g., Ranchhod and Zhou, 2001; Deutskens et al., 2004). Finding a way to increase the response rate would be highly beneficial. In some applications such as ad-hoc questionnaires regarding recent events, researchers aim for a fast response speed. Moreover, how well does the trigger-based survey tool, polling participants who are stating that they have new information perform? Second, previous work shows that in prediction markets experts can be identify ex post by their performance. This raises the question of when to ask the experts? This study tries to address the question of how to best acquire information from the market's experts. Additionally, online surveys are sovereign in response speed and the ability to methodologically control the filling process. According to Evans and Mathur (2005), online surveys fit best for the given purpose. Hence, a comparison of different types of online surveys is made. Specifically, common standalone surveys and market integrated surveys are used. The main research question, as introduced in Section 1.2 as research question 7 is:

> **Research Question 7:** *Are integrated surveys more accepted by participants of a prediction market than standalone surveys?*

The 'acceptance' is measured by response rate, total number of answered questions, and response speed. Additionally, a first experimental study on trigger-based surveys is run.

9.4 Results

In this section results of the study are presented. First, descriptive statistics are shown. Second, differences in response rate and reaction time are reported, followed by a closer look on the number of answered items. Third, an application of the trigger-based survey is presented and its response rate and acceptance is reported.

9.4.1 Descriptive Statistics

Participants were invited to fill out either the *integrated* or the *standalone* survey from 2013-07-29 until 2013-08-25 (28 days). Both groups are nearly equally sized (integrated: $N = 1,864$, $N_{active} = 706$; standalone: $N = 1,731$, $N_{active} = 657$) and hardly different in their trading activity (integrated: 71.57/7/466.83, standalone: 101.78/7/685.36 (#orders mean/median/sd)). Traders are counted as active, if they submitted at least one order while the market was active.

9.4.2 Response Rate and Reaction Time

The integrated survey has a higher response rate; both complete responses and partial responses are higher for participants using the integrated survey. The standalone survey leads to 6 (.91 %) complete and 38 (5.78 %) partial responses in contrast to 32 (4.53 %) complete and 124 (17.56 %) partial responses in the integrated survey. (Percentages relate to active traders in corresponding groups.) The advantage of integrated surveys over standalone surveys cannot conclusively shown here, due to the small response rates. Here, the integrated survey leads to an increase of 533.33 % in complete responses, and 326.32 % in partial responses compared to the standalone survey.

Next, the duration from the moment the survey was available until a participant answered his last question ('reaction time') is compared. In treatment *integrated*, participants that only *partly* answered the questionnaire have a median reaction time of 10.01 days (mean = 12.21) compared to 13.26 days (mean = 14.66) in treatment *standalone*. To completely fill the survey participants' reaction time is on median 6.52 days (mean = 7.80) in treatment *integrated* and 13.64 days (mean = 12.48) in treatment *standalone*. Summing up, treatment *integrated* delivers completely filled questionnaires (t-stat = 1.535, p = 6.68 %) as well as partially filled questionnaires significantly faster (t-value = 1.786, p = 3.94 %).

Result 9: *The integrated survey delivers results significantly faster (24.51 % – 52.20 %).*

9.4.3 Number of Answered Items

As shown, the difference between reaction times for *partial* responses is lower than for *complete* responses. As a rather long questionnaire was used, the number of *items* answered *in total* and *per survey participant* are also of interest. Those measures can help to estimate participants' acceptance of the duration of this particular survey and the possible response rate a shorter survey would have (had). In treatment *integrated* participants answered 3,522 items and 28.40 items per participant (median = 11). Treatment *standalone* leads to 2,497 answered items and 65.71 per participant (median = 64.71). Altogether, participants in treatment *integrated* answered on average significantly (Wilcoxon rank sum test, $W = 1,214$, $p < 0.01\,\%$) less items than in treatment *standalone*. It is assumed, that the reason lies in the 'entry barrier' of standalone surveys; i. e., in contrast to an integrated survey, which a participant might start, suspend, and continue as he pleases, an invitation link to a standalone survey represents a certain barrier. Participants might think twice before leaving the main web site to take part in a survey of unknown length and cognitive effort.

Result 10: *The integrated survey lead to 41.05 % more answered items.*

9.4.4 Trigger-based Survey

Last, the response rate for the trigger-based survey is analyzed. A total of 3,691 questions were triggered to 699 different traders resulting in 3,388 responses from 681 traders. As it cannot be distinguished if a participant answered "other reason" or denied to answer, the resulting response rate of 91.79 % is a lower bound. Participants decided to not answer a trigger-based survey 303 times (152 unique traders). 547 participants or 78.25 % never declined to answer a trigger-based survey.

Result 11: *Trigger-based survey: high response rate (91.79 %) and widely accepted (78.25 %).*

In order to illustrate how the two survey types can be used in conjunction, some preliminary data will be shown, matching both surveys on a per participant basis. Results of the trigger-based survey are shown in Table 9.1. Column "Answer" lists the possible responses as described in Section 9.3; column "Total" contains all data gathered with the trigger-based survey; column "Party/Candidate" shows only results of the trigger-based survey for traders who stated their preferred candidate/party via the integrated survey.

Obviously, answer 1 was the major reason for trades (66.59 %), regardless if traders submitted an order for their preferred party/candidate (62.04 %) or for any other party/candidate (65.35 %). Hence, the majority of orders were reportedly submitted based on (subjectively rational) economic considerations and thus might be more 'sensibly priced' as orders

TABLE 9.1: *Judgment Bias per Party*

	Total		Party/Candidate			
			Preferred		Other	
Answer	**Count**	**Percentage**	**Count**	**Percentage**	**Count**	**Percentage**
1) under/overvaluation	2,458	66.59 %	206	62.05 %	1,107	65.35 %
2) potential support/vote	511	13.84 %	75	22.59 %	267	15.76 %
3) like to see lower/higher price	175	4.74 %	22	6.63 %	76	4.49 %
4) bundle	244	6.61 %	18	5.42 %	112	6.61 %
5) other reason/no answer	303	8.20 %	11	3.31 %	132	7.79 %
Sum (answered)	3,388	91.79 %	321	96.69 %	1,562	92.21 %
Sum (total)	3,691	100.00 %	332	100.00 %	1,694	100.00 %

that were submitted for reasons 2 and 3. Similar to answer 1, answer 4 indicates subjectively rational trading behavior, as trading bundles is profit neutral. Altogether, over 80 % of participants reported to make their decisions based on rational considerations, which seems pretty reasonable. Although, there is a considerable proportion (18.58 %) where participants report, they want to see higher/lower prices or (not) support a certain candidate/party. At first glance, this seems to indicate irrational trading behavior. Nevertheless, this answer does not strictly exclude rational consideration (e. g., one might 'know' a certain candidate is overvalued, but prefer answer 3 over 1 anyway.). Therefore there is no reason to doubt participants answered truthfully on a large scale.

When comparing the percentual responses of all participants with those of participants that reported their party/candidate preferences, two salient contrasts can be seen. First, the differences between "Total" and "Preferred" are on average higher than for "Total" and "Other". Second, the biggest difference is present between "Total" and "Preferred" for answer 2 (8.75 %). Both observations might indicate, that participants' political preferences do affect their trading decisions. This might complement findings like the one that traders tend to buy more stocks of their preferred party (e. g., Kranz et al., 2014). Nevertheless, most differences between "Total" and "Preferred"/"Other" are rather small and – at first glance – the major tendency that roughly 2/3 chose answer 1 looks consistent.

Scratching the surface, it seems that trading behavior matches stated behavior in both surveys and survey responses are most widely consistent between both surveys. Summing up, the combination of integrated and trigger-based surveys provides a promising way to analyze individual trading behavior more deeply in further research.

Result 12: *The integrated survey and the trigger-based survey seem to deliver consistent results.*

9.5 Conclusion

In the age of near-ubiquitous internet access through an expanding variety of connected devices, it has not only become possible to conduct a greater number of polls, surveys, and preference elicitation tasks that involve a greater number of people; it is also possible to obtain more and richer information from each respondent. Nowadays, a popular way to gather that information on a continuous and repeated way is to run prediction markets. Their track record suggests that these markets may help to better foresee future developments and trends. Markets are powerful instruments for aggregating dispersed information, yet there are flaws. Markets are too complex for some users, they fail to capture massive amounts of their users' relevant information, and they suffer from some individual-level biases (e. g., Wolfers and Zitzewitz, 2004, 2006).

In this study a large-scale prediction market is integrated with a survey in two distinct ways. First, trading in prediction markets indicates that participants believe to have additional information. Consequently, a one-question survey is randomly triggered after a trade. The response rate for these types of question is with 91.79 % extremely high compared to typical online surveys. This approach reveals two advantages: (i) those participants who trade have information and (ii) they are actually interested to share that information. Both, having information and willingness to share information are usually out of scope of an online survey. Second, participants might find it convenient to answer survey questions right on the same platform. Hence, a survey feature was integrated. In a large-scale field experiment with over 3,500 participants, consistency and the response rate of the integrated survey was tested and compared to a standalone version. The integrated survey was found to deliver robust responses and a 3.16 % (533.33 % relative) higher response rate. Although the higher response rate could not be proven conclusively, both findings highlight the great possibilities for surveys to integrate with prediction markets. However, one has to be aware of structural differences between those survey types (e. g., participants might change their mind during an integrated survey).

Online surveys allow for qualitative responses and more complex question design. Especially given the fact that simple fill-in-the-blank and multiple choice questions give way to enhanced graphical interfaces that can capture probability distributions over response categories, even from people not familiar with distributions (Goldstein, 2013). Hence, combined forecasting with surveys and prediction markets can handle both; continuous, incentive-compatible forecasting as well as complex, quantitative question design.

As a direction for future research, it seems fruitful to develop adaptive survey systems that provide to ask participants only if the participant is expected to respond to it. In order to achieve that, it has to be evaluated whether it is possible to model the likelihood of a response. Moreover the trigger-based question, in its present form leads participants to rationalize their trading behavior. This leads to the question whether the trigger-based questionnaire leads to a different trading behavior. The limitations of the present work are straightforward: Most importantly, one instance of an integrated system of surveys and markets was explored in a political context. In order to increase external validity, the next step is to explore other information settings and implement surveys in prediction markets with other topics and populations.

Part IV

Finale

Chapter 10

Conclusion and Future Research

> “If you try to make something just to fit your uninformed view of some hypothetical market, you will fail. If you make something special and powerful and honest and true, you will succeed.”
>
> HUGH MACLEOD, 2004

10.1 Contributions

CONTINUOUS Market Engineering has shown to be suitable to create, maintain, and refine successful and sustainable markets that are able to adapt to changing demands and requirements. The work at hand documents the application of Continuous Market Engineering on two electronic markets. Applying the presented Continuous Market Engineering Process to both prediction markets studied in this work has lead to numerous improvements and insights. These contributions are demonstrated in the previous chapters, whereas this section recapitulates them in a more brief manner.

This thesis attempts to answer the following seven research questions:

- How do selected personal attributes (RA, CRA, and ERS) influence trading behavior in markets?

- How do selected personal attributes (RA, CRA, and ERS) influence decision quality in markets?
- How well can an unobtrusive analysis of trading behavior reveal trader preferences?
- Are decision behavior and decision outcome affected by the kind of device used?
- Is providing information about the disposition effect suitable to lower the strength of the disposition effect exhibited by an individual?
- Does a trend indicator arrow affect the strength of the disposition effect exhibited by an individual?
- Are integrated surveys more accepted by participants of a prediction market than standalone surveys?

The answers to these research questions can be summarized as follows.

Contribution 1: *How do selected personal attributes (RA, CRA, and ERS) influence trading behavior in markets?*
It has been known for long that personal attributes influence individual behavior. Nevertheless, the question of which attributes influence behavior in the domain of electronic markets is not completely answered yet. The work at hand specifically attempts to understand the interplay of risk aversion, cognitive reflection abilities, and emotion regulation strategies with two aspects of trading behavior, namely trading activity and liquidity providing. By analyzing trading data in conjunction with a questionnaire containing established personality tests, valuable insight on that interplay could be gained. Both cognitive reflection abilities (CRA) and risk aversion (RA) significantly improve trading activity (around 1.65 (RA) and 1.92 (CRA) additional orders per day). However, the tendency to provide liquidity is decreased by cognitive reflection and increased by risk aversion. Neither emotion regulation strategies (ERS) have shown to significantly influence trading activity, but both strategies have a significant impact on the tendency to provide liquidity. Participants using the emotion regulation strategy of suppression tend to provide liquidity, whereas participants using the reappraisal strategy tend to be liquidity takers. In summary, with only one exception, all regarded aspects of individual behavior are influenced by the assessed personal attributes. Specifically, only individual trading activity was not influenced by the emotion regulation strategy used.

Contribution 2: *How do selected personal attributes (RA, CRA, and ERS) influence decision quality in markets?*
Turning to personal attributes and decision quality, the following results can be summarized. First, cognitive abilities are shown to significantly improve trading performance (around P€ 120 per trade) and probability to make a profit. Second, risk aversion impairs trading performance (around P€ 90 per trade) as well as the probability to trade profitably. In contrast, risk-averse traders tend to provide liquidity. Third, emotion regulation strategies influence the regarded constructs. Emotion-suppressing traders perform significantly better (around P€ 270 per trade) and have a higher chance of submitting profitable orders, whereas using the reappraisal strategy leads to lesser gains (around P€ 100 per trade) and a smaller chance to decide profitably. Wrapping up, cognitive reflection and using the suppression strategy turns out to be beneficial for traders, whereas reappraisal and risk aversion impairs the regarded measures.

Contribution 3: *How well can an unobtrusive analysis of trading behavior reveal trader preferences?*
When predicting future events, it can be important to know whether and to which extent prediction market participants – in this case traders – stay objective or whether they are biased by their own preferences. Through a portfolio level analysis of trading data matched with survey data it is possible to consistently predict voters' intention in our market population. Furthermore, evidence is provided for a judgment bias consistent over all parties but elevated for small parties. Surprisingly, no significant difference occurs when comparing the bias between subgroups of tactical and non-tactical voters. Due to the bias' consistency across subgroups of tactical and non-tactical voters as well as different party preferences, a straightforward prediction model can be provided, which infers a trader's party preference from his trading behavior. The derived model correctly predicts party preferences in the validation set with an accuracy of around 70 %. Simply speaking, traders presumably unknowingly and unwillingly reveal themselves by their trading behavior. On the one hand, this implication somehow contradicts the traditional view on prediction markets, as anonymity is often seen as a major reason for participants to truthfully reveal their expectations – and thus eventually prediction markets' forecasting performance. Albeit, the intra-participant anonymity stays unaffected, awareness of this result might reduce participants' *perceived* anonymity and thus their trust attitude towards prediction markets (especially when dealing with 'sensitive' questions). This can in turn impair participants' willingness to participate on such markets. On the other hand, it supports existing theory on prediction markets by showing that traders tend to reveal their true preferences

on prediction markets by submitting orders accordingly. In the end, preference-consistent trading behavior – as encountered in this prediction market – is the foundation to infer individuals' preferences from their trading behavior.

Contribution 4: *Are decision behavior and decision outcome affected by the kind of device used?*
After demonstrating the interplay of personal attributes and preferences with decision outcome, the influence of device choice was shown. In a comparison of mobile and stationary trading interfaces it is illustrated that orders submitted from a mobile device perform significantly worse than orders submitted via a stationary trading interface. Interestingly, a significant influence of the device choice on the traders confidence (measured with two proxies) could not be proven. This result illustrates once more that decision making performance does not solely depend on the decision maker and his (cognitive) resources. Instead, awareness of differences between devices when designing software artifacts has shown to be important.

Contribution 5: *Is providing information about the disposition effect suitable to lower the strength of the disposition effect exhibited by an individual?*
A well-explored deviation from rational behavior is the disposition effect. Its negative impact on trading performance has been shown in many studies of trading environments. To assess the individual and aggregated disposition effect, a field study was conducted on a prediction market on which the disposition effect could be determined for over 500 traders. In line with previous research, a significant disposition effect on both the aggregated and the individual level was found in the regarded market. As a debiasing method, half the participants could access an information text about the disposition effect. By analyzing the information text access, it could not be shown that this textual information is suitable to significantly impact the traders' individual disposition effect.

Contribution 6: *Does a trend indicator arrow affect the strength of the disposition effect exhibited by an individual?*
Although creating awareness of the disposition effect via textual information could not verifiably decrease its strength, even tiny visual cues have shown to significantly increase the disposition effect's strength and thus letting traders deviate from rational behavior. Specifically, a performance indicator arrow transparently showing if a trader's portfolio value is positive or not significantly increases the disposition effect by around 12 % compared to the market's average (up to 71 % compared between treatment groups). This result is of particular importance, as such interface elements are widely used in online brokerage in-

terfaces, and thus also has an implication for regulators. Hence, it might be worthwhile to consider mandating online brokers to educate customers about behavioral biases, which are known to degrade individual performance. As the previous result (cf. Contribution 5) implies, textual advice seems not to be a proper solution in this case. Thus, it might be sensible to consider interface design itself as a regulatory object.

Contribution 7: *Are integrated surveys more accepted by participants of a prediction market than standalone surveys?*
In the age of near-ubiquitous internet access through an expanding variety of connected devices, it has not only become possible to conduct a greater number of polls, surveys, and preference elicitation tasks that involve a larger number of people; it is also possible to obtain more and richer information from each respondent. In this study, two types of surveys were combined with a prediction market on which over 1,400 participants traded actively. Following the assumption that conducting a trade indicates that a participant believes to have information about the product traded, a one-question survey ('post-trade question') was randomly triggered consequently upon a trade. Compared to typical response rates of online surveys, this post-trade question's response rate was with 91.79 % extremely high. Assuming that participants of an online platform prefer to answer surveys right on the same platform due to convenience reasons, a second study was conducted. It aimed to identify if the effort to integrate a rather long survey on an existing platform – in contrast to using a specialized standalone survey system – pays out through a substantial increase in response rate or data quality. Hence, half of the participants were offered an integrated survey system, whilst the other half were redirected to an external survey system. Subsequently, the response rate and consistency of both surveys were compared. The integrated survey delivered robust responses and a 3.16 % (533.33 % relative) higher response rate. Albeit the higher response rate could not be proven conclusively, these results emphasize the potential in integrating surveys into online markets.

All presented contributions were derived from studies carried out on a special form of web-based electronic markets, namely prediction markets. This implies the presence of certain market specifics (play money, virtual products, and the like) most other electronic markets lack. Furthermore, the conducted studies mainly focus on the three facets *Agent Behavior*, *Interfaces*, and *Auxiliary Services*. Consequently, the unrestricted generalizability of the results and their implications for any market might not be given. Nevertheless, the reported contributions can be seen as a first step in demonstrating how to refine existing electronic markets against the backdrop of Continuous Market Engineering. Thus, these studies pave the way for market engineers in identifying entry points for improving markets.

10.2 Outlook

Having presented the main contributions from applying Continuous Market Engineering on two Prediction Markets, on-going work and ideas for future research are described.

Improve questionnaire on trader's Market Predisposition The presented set of personal attributes assumed to influence individual trading behavior and thus representing an individual's Market Predisposition is – albeit suitable – neither complete nor optimized (cf. Chapter 5.5). Further research is needed to improve the questionnaire used. First, it seems fruitful to evaluate additional constructs that might have an explanatory value for a trader's Market Predisposition. Second, especially since it most certainly will be beneficial to include additional constructs, the duration of the Market Predisposition test is object to optimization. One starting point can be the use of reduced alternatives to established tests, e. g., a shortened version of the Ten-Paired Lottery by Holt and Laury (2002) as presented in Teubner (2013) or the application of a different risk preference test (see, e. g., Charnes et al. (2013) for an overview).

Feedback on trader's Market Predisposition and adaptive interfaces A straightforward question on assessing a trader's Market Predisposition is whether traders that are informed about their individual Market Predisposition (i) would alter their decision to join a certain market and (ii) if they will change their willingness to improve their own abilities according to their measured predisposition immediately or over the course of time. Those questions could be answered in a follow-up study on a given electronic market with minor intervention. To examine question (i), existing and potential participants would have to be separated in one control as well as one treatment group. In the treatment group each existing participant would have to undergo a yet to be refined Market Predisposition test. Potential participants in the treatment group would have to conduct the very same test as first step of the market's registration process. The result of the test would then have to be presented to the participants instantly and individually. An accompanying survey could be used in the proposed study in oder to reflect the participants' understanding of the presented results and their implications. In order to examine question (ii), additional panel surveys would have to be conducted amongst control and treatment groups in a yet to be defined interval. For a market provider, like an online retail brokerage company, such an individual measure could be used to easily cluster participants along their predisposition and thus their inherent need for support. Practically, it might be beneficial for participants with a lower predisposition to be redirected to a more simplified and supportive trading interface whereas the more proficient traders can be confronted with a full-featured expert interface.

Comparing Web and Mobile interface in a laboratory experiment The study presented in Chapter 7 showed that even though participants' decision confidence was not decreased, when participants chose to submit orders via a mobile device their decision performance suffered. As discussed in Section 7.5 the conducted field study is by design not suited to examine questions of environmental influences and participant's acceptance. To rule out that the performance decrease is solely driven by environmental factors, a laboratory experiment could be conducted. Such an experiment offers the chance to gain additional insights; for instance concerning acceptance of the mobile interface offered via a Technology Acceptance study (cf. Venkatesh, 2000; Venkatesh and Davis, 2000).

Control for understanding of Disposition Effect information text As has been shown in Section 8.4.1, only about one quarter of participants accessed the offered disposition effect information text. Even for participants accessing that information text, it was neither possible to validate that (i) the text and its implications were understood nor (ii) read at all. A laboratory experiment could be set up to control for these factors; an additional questionnaire could provide certainty if a participant has read and understood the concept of the disposition effect and its implications on trading performance. Furthermore, in such a follow-up laboratory experiment, a reposition of the offered link in a more conspicuous location is worth considering.

Validate superiority of integrated surveys The integrated survey in its present form (see Section 9.3) proved its superiority over an external specialized survey system in one specific market instance. The next logical step is to increase external validity. Hence, it seems fruitful to explore other settings of online communities and different lengths and types of surveys as well as different survey systems.

Towards a combination of Prediction Markets and Delphi Method Accompanied by the motivating results of the survey integration study (see Chapter 9), a combination of Prediction Markets with a more sophisticated kind of survey technique, namely the Delphi Method (cf. Dalkey and Helmer, 1963), appears encouraging to further improve predictive quality by gathering additional information from participants for the following reason. On the one hand, Prediction Markets are known to be very apt performing quantitative predictions and are able to react rapidly to changing situations and newly emerged information. However, they are not a very appropriate forecasting instrument for quantitative or conditional predictions – albeit it is possible to apply them to such settings. On the other hand, the Delphi Method is profoundly suited for such kind of settings. Nevertheless, it has a drawback when it comes to selecting the 'right' participants in sense of information or expertise of subjects, which in turn is a major strength of prediction markets, as well performing (and thus, in a way, predicting) participants can be easily identified by their past performance (cf. Equation 7.1). The dynamics of a Prediction Market offers additional opportunities to identify situations in which participants might have important information. An indication for such a situation is, for instance, when a trader that used to predict well in the past submits an order with a limit price that strongly diverges from the current market price. In such a case, the presented trigger-based question (Section 9.3.1, Figure 9.2b) can be used to instantly prompt that participant for the reason behind his action. As shown in Chapter 9, the technical foundations for integrated surveys have already been laid as part of the work at hand in case of the Economic Indicator eXchange. However, important questions on how to combine those forecasting approaches remain open for further research.

Examine rationalization effect of trigger-based question The trigger-based question (see Section 9.3.1) leads participants to rationalize their trading behavior. This is achieved by simply prompting the market participant immediately after submitting an order to answer why the order was submitted with respect to the current market price of the product in question and the order's limit price. Answering such a question forces the participant to rationalize ones behavior retrospectively, which in turn leads to the question whether the trigger-based questionnaire leads to a different trading behavior in subsequent trading decisions. At first glance, present data seems to support the hypothesis that traders have a tendency to decide more on rational reasons after being confronted to a trigger-based question. Nevertheless, further research is required to answer that question adequately.

Enhance Prediction Markets with User Generated Content Prediction markets make use of the so-called *Wisdom of Crowds* (cf. Surowiecki, 2004) to obtain their predictive power. The participating actors of an actual prediction market instance can be separated in market participants (*Agents*) and market operators (*Market Engineers*). Currently, most prediction market providers are responsible for the technical market operation, a certain set of trading rules and the selection of forecasting goals (i. e., tradable products). Whereas the market participants' stake is to trade their information on that market and thereby generating and improving a certain prediction. A more progressive approach could further enable the *Crowd* to participate in the market, not only by predicting, but instead by contributing *User-Generated Content* (cf. Krumm et al., 2008; van Dijck, 2009) through defining the forecasting goals itself. Ultimately, such an enhanced prediction market system would shift the market providers role to (i) a pure provider of a trading rule set and (ii) a technical market operator, whereas the participants role would broaden towards creation of content in the form of products. This flexibility would open up a broad range of new questions regarding interface design and user guidance in setting up new product proposals. In order to prevent thin markets, a major challenge will be to figure out a mechanism for deciding which product proposal to approve for trading. One approach would surely be to define an *Initial Public Offering* (IPO) phase in which the product proposal is reviewed by the market provider, presented to the public, and has to raise a certain amount of money (cf. *Book Building*) in order to get approval for future trading. In case a defined amount of money can be raised, one can assume that sufficient interest of participants exists to trade the proposed product. The evaluation of the illustrated approach could take place on an existing prediction market, since by design, user-generated products would be able to co-exist with 'traditional' products.

Incentive Engineering As discussed in Section 3.3, a properly designed incentive system is essential for a play-money prediction market. Straub, Gimpel, and Teschner (2014) showed that incentive systems influence the effort so-called *crowd workers* are willing to invest for a certain task. In a related study Straub, Gimpel, Teschner, and Weinhardt (2014) investigate the influence of competitors' performance on individuals' effort (i. e., task completion rates) as well as on individuals' premature abandonment rates. They find that the strength of competitors decreases both regarded individual measures. It seems reasonable to believe that those relationships also tend to hold in the domain of prediction markets. A future study might examine, (i) whether those results are applicable to prediction markets and – if so – (ii) whether and which additional aspects of their trading behavior are influenceable by incentive design decisions. Ultimately, guidelines for a purposefully *Incentive Engineering* might be derived from this research.

10.3 Summary

In this Chapter, the main contributions to literature based on the results contained in this thesis were presented. First, a Continuous Market Engineering Process based on related theory was presented and discussed. Subsequently, the application of Continuous Market Engineering methodology on two play-money prediction markets with a focus on Agent Behavior, Interfaces, and Auxiliary Services was documented. The results presented help market engineers to better understand trader behavior in electronic markets and highlight numerous approaches to analyze, support, and improve traders' market interaction. Finally, several opportunities for further research focusing on Agent Behavior in the context of Continuous Market Engineering have been sketched. The outlined research directions are appropriate to further explore participants' behavior in electronic market systems. I would be pleased if the contributions of the work at hand help to widen the understanding of dynamics in markets and inspiring further research in this thrilling research domain.

Part V

Appendices

Appendix A

Texts and Trade Screens

Disposition Effect Info Text

Information text used in the study presented in Chapter 8.

Original:

> "Entgegen der alten Börsenregel, nach der Anleger Gewinne laufen lassen und Verluste durch Verkauf begrenzen sollen, ist an fast allen internationalen Kapitalmärkten genau das Gegenteil zu beobachten. Aufgelaufene Gewinne werden meist zu früh realisiert, während sich Anleger von Aktien in der Verlustzone häufig viel zu spät trennen. Diese irrationale Verhaltensweise wird als Dispositionseffekt bezeichnet. Der Dispositionseffekt führt häufig zu individuellen Verlusten und entgangenen Gewinnen. Anleger, die den Dispositionseffekt vermeiden, können deutlich höhere Gewinne erzielen."

Author's translation:

> "In contrast to the well-known stock market rule that investors should ride gains and sell losses, precisely the opposite is observable from nearly all international capital markets. Profitable stocks are often sold too early, whilst traders depart from losing stocks way too late. This irrational behavior is known as disposition effect. The disposition effect often leads to trader losses and missed gains. Investors, who resist the disposition effect may realize significantly higher gains."

Trade Screens per Treatment

This Section contains screenshots for the different treatments presented in Chapter 8. For an overview refer to Table A.1.

TABLE A.1: *Overview of Treatments and Screenshots*

	DE Info Text	**w/o DE Info Text**
Trend Indicator	*Trend_Info* (Figure 8.1)	*Trend* (Figure A.1)
w/o Trend Indicator	*Info* (Figure A.2)	*Control* (Figure A.3)

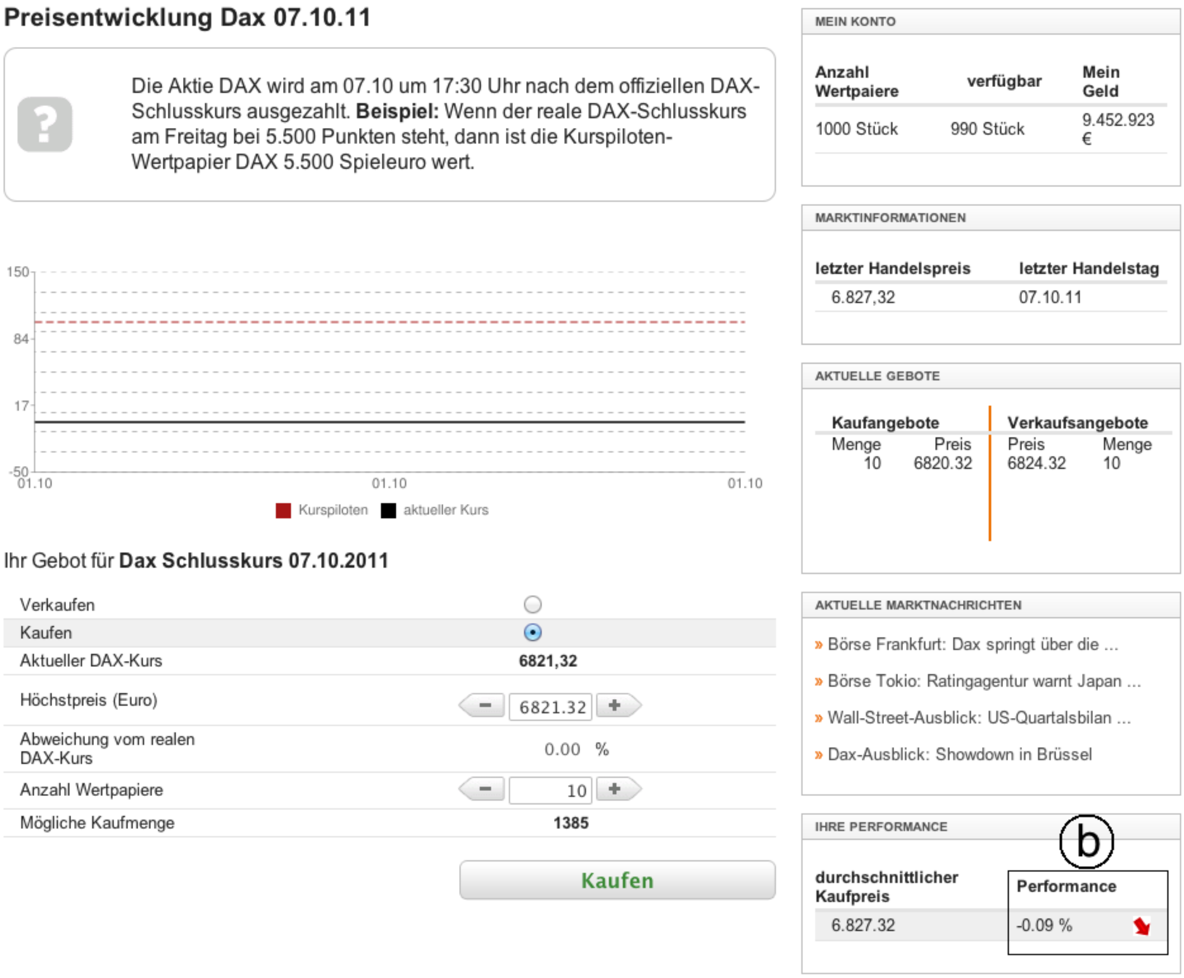


FIGURE A.1: *Trading Screen for Treatment Trend*

(Containing one user interface modification. Modification (b) shows the 'Trend Indicator' element.)

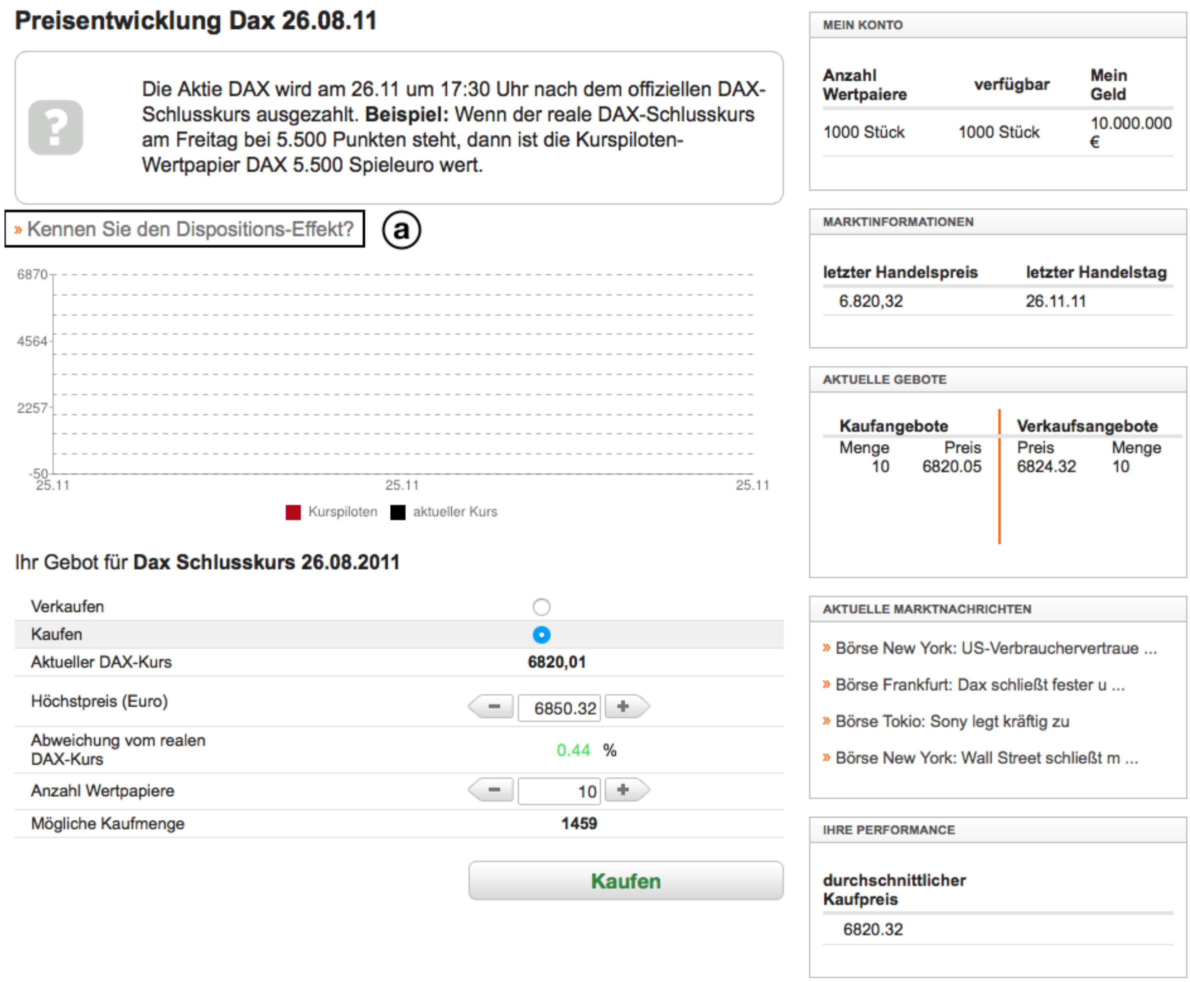


FIGURE A.2: *Trading Screen for Treatment Info*

(Containing one user interface modification. A click on the linked text (a) fades in an info text about the disposition effect. The whole text is depicted in Appendix A.)

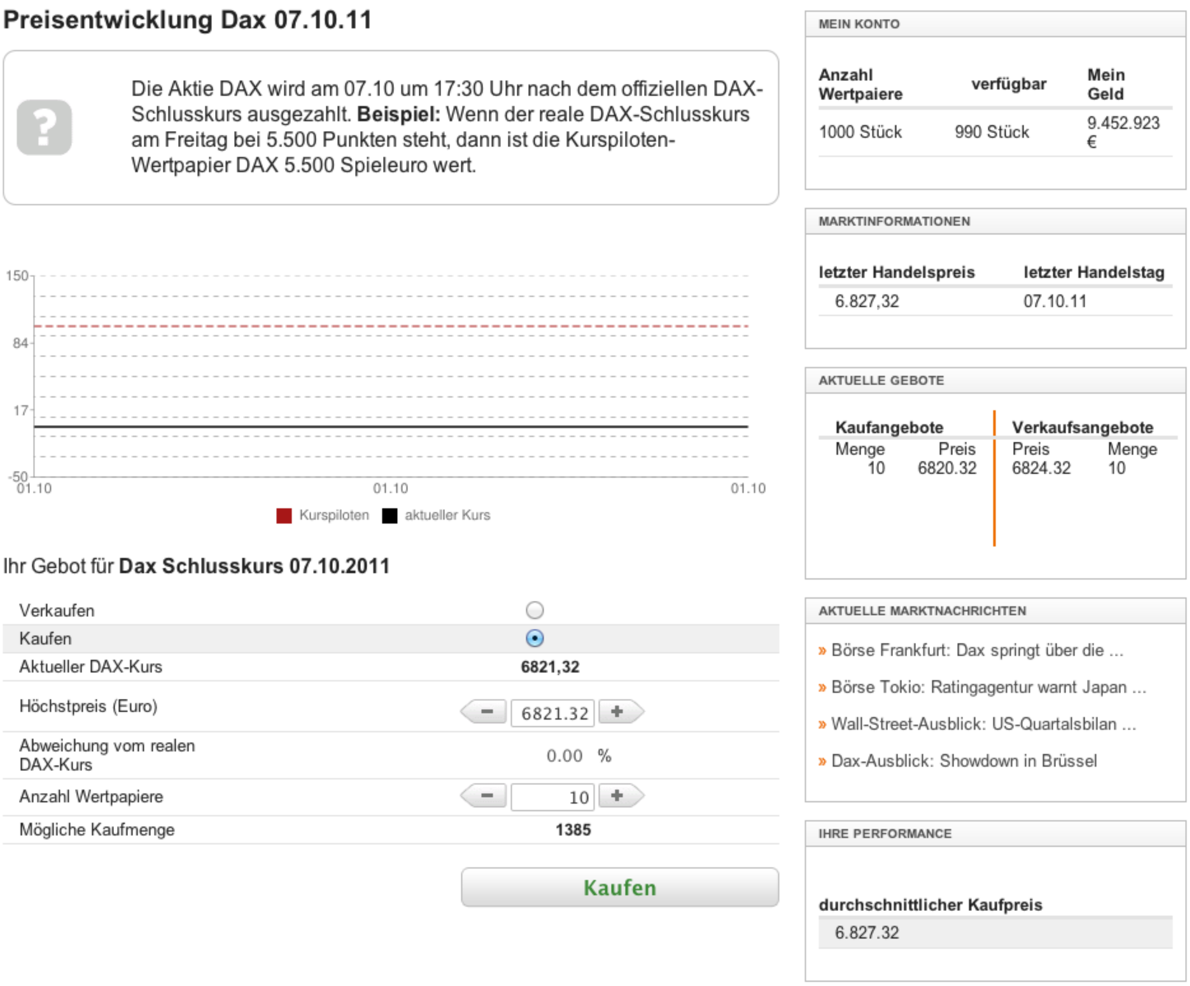


FIGURE A.3: *Trading Screen for Treatment Control*

Appendix B

Heath et al. measure

The adapted Heath et al. measure, as used in the study presented in Chapter 6 is shown in Table B.2. The original Heath et al. measure is depicted in Table B.1.

TABLE B.1: *Wording of the Heath et al. measure (Heath et al., 1991)*

	Which of these comes closest to the reason for your decision with regard to the [. . .] in the last federal election?	
[. . .]	**N**	**Reply**
first vote	1	I always vote that way.
	2	I thought it was the best candidate.
	3	I really preferred another candidate, but he had no chance of winning in this constituency.
	4	Other reason
	5	Prefer not to say
second vote	1	I always vote that way.
	2	I thought it was the best party.
	3	I really preferred another party, but I wanted to allow another party to overcome the 5% threshold.
	4	Other reason
	5	Prefer not to say

TABLE B.2: *Wording of the modified Heath et al. measure (based on Heath et al., 1991)*

	Welche dieser Aussagen lag Ihrer Wahlentscheidung zur [...] in der letzten Bundestagswahl am ehesten zu Grunde?	
[...]	**N**	**Reply**
Erststimme	1	Ich wähle immer so.
	2	Aus meiner Sicht war es der beste Kandidat.
	3	Ich zog einen anderen Kandidaten vor, aber dieser hatte keine Aussicht auf den Wahlsieg.
	4	anderer Grund
	5	keine Angabe
Zweitstimme	1	Ich wähle immer so.
	2	Aus meiner Sicht war es die beste Partei.
	3	Ich zog eine andere Partei vor, wollte aber einer weiteren Partei das Überwinden der 5%-Hürde ermöglichen.
	4	anderer Grund
	5	keine Angabe

References

Adipat, B. and D. Zhang (2005, August). Adaptive and personalized interfaces for mobile web. In *15th Annual Workshop on Information Technologies & Systems (WITS)*, pp. 1 – 6.

Al-Ubaydli, O. and P. Boettke (2011, May). Markets as economizers of information: Field experimental examination of the "hayek hypothesis". *GMU Working Paper in Economics* (11-10), 1 – 14.

Ammon, U. (2009). *Delphi-Befragung, Handbuch Methoden der Organisationsforschung. Quantitative und Qualitative Methoden*. VS Verlag, Wiesbaden.

Anand, P. and B. Sternthal (1989). *Cognitive and Affective Responses to Advertising*, Chapter Strategies for designing persuasive messages: Deductions from the resource matching hypothesis, pp. 135 – 159. England: Lexington.

Andreassen, P. B. (1988, June). Explaining the price-volume relationship: The difference betweeen price changes and changing prices. *Organizational Behavior and Human Decision Processes 41*(3), 371 – 389.

Andres, L. (2012). *Designing & Doing Survey Research*. 1 Oliver's Yard, 55 City Road, London EC1Y 1 SP: SAGE Publications Ltd.

Ang, J. S. and T. Schwarz (1984, July). Risk aversion and information structure: An experimental study of price variability in the securities markets. *The Journal of Finance 40*(3), 825 – 844.

Antweiler, W. (2012). Long-term prediction markets. *The Journal of Prediction Markets 6*(3), 43 – 61.

Ariely, D. (2000). Controlling the information flow: Effects on consumers' decision making and preferences. *Journal of Consumer Research 27*, 233 – 248.

Arrow, K. J., R. Forsythe, M. Gorham, R. Hahn, R. Hanson, J. O. Ledyard, S. Levmore, R. Litan, P. Milgrom, F. D. Nelson, G. R. Neumann, M. Ottaviani, T. C. Schelling, R. J. Shiller, V. L. Smith, E. Snowberg, C. R. Sunstein, P. C. Tetlock, P. E. Tetlock, H. R. Varian,

J. Wolfers, and E. Zitzewitz (2008, May). The promise of prediction markets. *Science 320*, 877 – 878.

Ash, S. (2007, October). MoSCoW Priorisation. Technical report, DSDM Consortium.

Babad, E., M. Hills, and M. O'Driscol (1992). Factors influencing wishful thinking and predictions of election outcomes. *Basic and Applied Social Psychology 13*(4), 461 – 476.

Babad, E. and Y. Katz (1991). Wishful thinking – against all odds. *Journal of Applied Social Psychology 21*(23), 1921 – 1983.

Babad, E. and E. Yacobos (1993, March). Wish and reality in voters' predictions of election outcomes. *Political Psychology 14*(1), 37 – 54.

Babbie, E. (1990). *Survey Research Methods* (Second ed.). 20 Davis Drive, Belmont, CA 94002, USA: Wadsworth.

Bakos, Y., H. C. J. Lucas, and W. Oh (2005, December). The impact of e-commerce on competition in the retail brokerage industry. *Information Systems Research 16*(4), 352 – 371.

Barber, B. M., Y.-T. Lee, Y.-J. Liu, and T. Odean (2007). Is the aggregate investor reluctant to realise losses? evidence from taiwan. *European Financial Management 13*(3), 423 – 447.

Barber, B. M. and T. Odean (2001, February). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics 116*(1), 261 – 292.

Beck, K. (2000). *Extreme Programming Explained*. Addison-Wesley Professional.

Beck, K. (2002, November). *Test Driven Development by Example*. Addison-Wesley Professional.

Belanoff, P., P. Elbow, and S. I. Fontaine (1991, January). *Nothing begins with N: New investigations of freewriting*. Southern Illinois University Press.

Bennouri, M., H. Gimpel, and J. Robert (2011). Measuring the impact of information aggregation mechanisms: An experimental investigation. *Journal of Economic Behavior & Organization 78*(3), 302 – 318.

Berg, J., R. Forsythe, F. Nelson, and T. Rietz (2008). Results from a dozen years of election futures markets research. In C. R. Plott and V. L. Smith (Eds.), *Handbook of Experimental*

Economics Results, Volume 1 of *Handbook of Experimental Economics Results*, pp. 742 – 751. Elsevier.

Berg, J. E., F. D. Nelson, and T. A. Rietz (2008). Prediction market accuracy in the long run. *International Journal of Forecasting 24*(2), 285 – 300.

Berg, J. E. and T. A. Rietz (2003). Prediction markets as decision support systems. *Information Systems Frontiers 5*, 79 – 93. 10.1023/A:1022002107255.

Berg, J. E. and T. A. Rietz (2010, July). Longshots, overconfidence and efficiency on the iowa electronic market. Available at SSRN: `http://ssrn.com/abstract=1645062`.

Berlemann, M. and C. Schmidt (2001). Predictive accuracy of political stock markets: Empirical evidence from a european perspective. *Discussion Papers, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes 57*, 1 – 42.

Berners-Lee, T., R. Fielding, and H. Frystyk (1996, May). Rfc 1945. http://www.ietf.org/rfc/rfc1945.txt.

Berners-Lee, T. and M. Fischetti (2000). *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor*. HarperInformation.

Bitkom (2013, May). Online-kampagnen entscheiden die bundestagswahl. Available at `http://www.bitkom.org/de/presse/8477_76080.aspx`.

Block, C. (2010). *Agile Market Engineering*. Dissertation, KIT-IISM.

Blume, M., S. Luckner, and C. Weinhardt (2010, November). Fraud detection in play-money prediction markets. *Information Systems and e-Business Management 8*(4), 395–413.

Boer, H., E. T. Huurne, and E. Taal (2006, June). Effects of pictures and textual arguments in sun protection public service announcements. *Cancer Detection and Prevention 30*(5), 432 – 438.

Boer, K., U. Kaymak, and J. Spiering (2007, May). From discrete-time models to continuous-time, asynchronous modelling of financial markets. *Computational Intelligence 23*(2), 142 – 161.

Borghesi, R. (2009). The effect of contract structure on prediction market price biases. *The Journal of Prediction Markets 3*(3), 1 – 12.

Borghesi, R. (2013, May). The impact of the disposition effect on asset prices: Insight from the nba. *Journal of Economics and Finance Forthcoming*, 1 – 14.

Box, G. E. and G. M. Jenkins (1976). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Brewster, S. (2002, May). Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous Computing 6*(3), 188 – 205.

Buzan, T. and B. Buzan (1993). *The mind map book: Radiant thinking – the major evolution in human thought*. London: BBC Publications.

Camargo, A. and I. Fonseca (2013, March). The race for self-directed investors: Developments in online trading among brokers and banks. Technical Report 1, Celent, 499 Washington Blvd, 11th Floor Jersey City, NJ 07310.

Charnes, G., U. Gneezy, and A. Imas (2013, December). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization 87*, 43 – 51.

Chen, Y., C.-H. Chu, T. Mullen, and D. M. Pennock (2005). Information markets vs. opinion pools: An empirical comparison. In *Proceedings of the 6th ACM conference on Electronic commerce*, pp. 58 – 67. ACM.

Chen, Y. and D. M. Pennock (2012). A utility framework for bounded-loss market makers. Papers, arXiv.org.

Christiansen, J. D. (2007). Prediction markets: Practical experiments in small markets and behaviour observed. *The Journal of Prediction Markets 1*(1), 17 – 41.

Clegg, D. and R. Barker (1994). *CASE Method Fast-Track: A RAD Approach*. Boston, MA, USA: Addison-Wesley Longman Publishing.

Cook, C., F. Heath, and R. L. Thompson (2000, December). A meta-analysis of response rates in web- or internet-based surveys. *Educational and Psychological Measurement 60*(6), 821 – 836.

Cooper, R. G. (1990, May). Stage-gate systems: A new tool for managing new products. *Business Horizons 33*(3), 44 – 54.

Cronbach, L. J. (1951, September). Coefficient alpha and the internal structure of tests. *Psychometrika 16*(3), 297 – 334.

Cuhls, K. (2003, April). From forecasting to foresight processes—new participative foresight activities in germany. *Journal of Forecasting 22*(2-3), 93 – 111.

Dalkey, N. and O. Helmer (1963). An experimental application of the delphi method to the use of experts. *Management Science 9*(3), 458 – 467.

Deutskens, E., K. D. Ruyter, M. Wetzels, and P. Ooserveld (2004). Response rate and response quality of internet-based surveys. *Marketing Letters 15*(1), 21 – 36.

Dillman, D. A. (1991). The design and administration of mail surveys. *Annual Review of Sociology 17*, 225 – 249.

Dittrich, D., W. Güth, and B. Maciejovsky (2005, February). Overconfidence in investment decisions: An experimental approach. *The European Journal of Finance 11*(6), 471 – 491.

Edelman, B., M. Ostrovsky, and M. Schwarz (2007, March). Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *The American Economic Review 97*(1), 242 – 259.

Erikson, R. S. and C. Wlezien (2008). Are political markets really superior to polls as election predictors? *Public Opinion Quarterly 72*(2), 190 – 215.

Eriksson, N. (2012a, January). A follow up of internet adopters' use, perceptions and channel preferences of electronic travel services. In M. Fuchs, F. Ricci, and L. Cantoni (Eds.), *Information and Communication Technologies in Tourism 2012*. Springer.

Eriksson, N. (2012b, April). User experience of trip arrangements: A comparison of mobile device and computer users. *International Journal of E-Services and Mobile Applications 4*(2), 55 – 69.

Evans, J. R. and A. Mathur (2005). The value of online surveys. *Internet Research 15*(2), 195 – 219.

Fama, E. F. (1965, January). The behavior of stock-market prices. *The Journal of Business 38*(1), 34 – 105.

Fama, E. F. (1970, May). Efficient capital markets: A review of theoretical and empirical work. *The Journal of Finance 25*(2), 383 – 417.

Fama, E. F. (1991). Efficient capital markets: Ii. *The Journal of Finance 46*(5), 1575 – 1617.

Fellner, G. and B. Maciejovsky (2007, June). Risk attitude and market behavior: Evidence from experimental asset markets. *Journal of Economic Psychology 28*(3), 338 – 350.

Feng, L. and M. S. Seasholes (2005). Do investor sophistication and trading experience eliminate behavioral biases in financial markets. *Review of Finance 9*, 305 – 351.

Fenton-O'Creevy, M., E. Soane, N. Nicholson, and P. Willman (2011, July). Thinking, feeling and deciding: The influence of emotions on the decision making and performance of traders. *Journal of Organizational Behavior 32*, 1044 – 1061.

Fielding, R., J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee (1999, June). Rfc 2616. http://www.ietf.org/rfc/rfc2616.txt.

Fisher, S. D. (2004, January). Definition and measurement of tactical voting: The role of rational choice. *British Journal of Political Science 34*(1), 152 – 166.

Fornahl, D. and M.-P. Menzel (2007). Cluster life cycles: Dimensions and rationales of cluster development. Jena Economic Research Papers 2007, 076, Jena.

Forsythe, R., F. Murray, V. Krishnamurthy, and T. W. Ross (1998). Markets as predictors of election outcomes: Campaign events and judgement bias in the 1993 ubc election stock market. *Canadian Public Policy 24*(3), 329 – 351.

Forsythe, R., F. Nelson, G. R. Neumann, and J. Wright (1992, December). Anatomy of an experimental political stock market. *The American Economic Review 82*(5), 1142 – 1161.

Forsythe, R., T. A. Rietz, and T. W. Ross (1999, May). Wishes, expectations and actions: a survey on price formation in election stock markets. *Journal of Economic Behavior & Organization 39*(1), 83 – 110.

Fowler, F. J. J. (2014). *Survey Research Methods* (Fifth ed.). SAGE Publications Ltd.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives 19*(4), 25 – 42.

Fromm, H., F. Habryn, and G. Satzger (2012). Service analytics: Leveraging data across enterprise boundaries for competitive advantage. In *Globalization of Professional Services*, pp. 139 – 149. Berlin Heidelberg: Springer.

Garvey, R. and A. Murphy (2004, July). Are professional traders too slow to realize their losses? *Financial Analysts Journal 60*(4), 35 – 43.

Gimpel, H., N. R. Jennings, G. E. Kersten, A. Ockenfels, and C. Weinhardt (2008). *Negotiation, Auctions, and Market Engineering*, Volume 2 of *Lecture Notes in Business Information Processing*, Chapter Market Engineering: A Research Agenda, pp. 1 – 15. Springer, Berlin.

Gjerstad, S. (2005, January). Risk aversion, beliefs, and prediction market equilibrium. Microeconomics, EconWPA.

Goel, S., D. M. Reeves, D. J. Watts, and D. M. Pennock (2010). Prediction without markets. In *Proceedings of the 11th ACM conference on Electronic commerce*, New York, NY, USA, pp. 357 – 366. ACM.

Goldstein, B. E. (2013). *Sensation and Perception*. Wadsworth.

Gordon, T. J. (2007). Energy forecasts using a "roundless" approach to running a delphi study. *Foresight 9*(2), 27 – 35.

Graefe, A., S. Luckner, and C. Weinhardt (2010). Prediction markets for foresight. *Futures 42*, 394 – 404.

Griffin, A. (1997, November). Pdma research on new product development practices: Updating trends and benchmarking best practices. *Journal of Product Innovation Management 14*(6), 429 – 458.

Gross, J. J. and O. P. John (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology 85*(2), 348 – 362.

Güth, W., J. P. Krahnen, and C. Rieck (1997, April). Financial markets with asymmetric information: A pilot study focusing on insider advantages. *Journal of Economic Psychology 18*(2-3), 253 – 257.

Hahn, R. W. and P. C. Tetlock (2005). Using information markets to improve public decision making. *Harvard Journal of Law & Public Policy 29*(1), 214 – 289.

Hammond, D. (2011, May). Cigarette package health warnings and interest in quit smoking – 14 countries. *Morbidity and Mortality Weekly Report 60*(20), 645 – 651.

Hanson, R. (2003). Combinatorial information market design. *Information Systems Frontiers 5*(1), 107 – 119.

Hanson, R. (2007). Logarithmic market scoring rules for modular combinatorial information aggregation. *The Journal of Prediction Markets 1*(1), 3 – 15.

Hanson, R. D. (1999). Decision markets. *IEEE Intelligent Systems 14*(3), 16 – 19.

Harris, L. (2003). *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford University Press.

Hartzmark, S. M. and D. H. Solomon (2012). Efficiency and the disposition effect in nfl prediction markets. *Quarterly Journal of Economics 2*(3), 1250013–1 – 1250013–42.

Heath, A. F., R. Jowell, J. Curtice, G. Evans, J. Field, and S. Witherspoon (1991). *Understanding political change: the British voter, 1964-1987*. Oxford: Pergamon Press.

Heilemann, U. and H. O. Stekler (2012, May). Has the accuracy of macroeconomic forecasts for germany improved? *German Economic Review 14*(2), 235 – 253.

Henderson-Sellers, B. and J. M. Edwards (1990). The object-oriented systems life cycle. *Communications of the ACM 33*(9), 142 – 159.

Hillygus, D. S. (2011). The evolution of election polling in the united states. *Public Opinion Quarterly 75*(5), 962 – 981.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American economic review 92*(5), 1644 – 1655.

Huber, J., M. Kirchler, and M. Sutter (2008, January). Is more information always better: Experimental financial markets with cumulative information. *Journal of Economic Behavior & Organization 65*(1), 86 – 104.

Ilieva, J., S. Baron, and N. M. Healey (2002). Online surveys in marketing research: Pros and cons. *International Journal of Market Research 44*(3), 361 – 376.

Issenberg, S. (2012, December). How president obama's campaign used big data to rally individual voters. Available at `http://www.technologyreview.com/featuredstory/509026/how-obamas-team-used-big-data-to-rally-voters/`.

Jacobsen, B., J. Potters, A. Schram, F. van Winden, and J. Wit (2000). (in)accuracy of a european political stock market: The influence of common value structures. *European Economic Review 44*(2), 205 – 230.

Jacoby, J. (1984, March). Perspectives on information overload. *Journal of Consumer Research 10*(4), 432 – 435.

Jäger-Ambrozewicz, M. (2009, November). Eix-wochenschau nr. 1. Available at `http://eix-market.de/static/images/pdf/Newsletter_Nr_1.pdf`.

Jeffries, R. (2001, August). Essential xp: Card, conversation, confirmation. Available at http://xprogramming.com/articles/expcardconversationconfirmation/.

Kahneman, D. and A. Tversky (1979, March). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263 – 292.

Kano, N., N. Seraku, F. Takahashi, and S. Tsuji (1984). Attractive quality and must-be quality. *Journal of the Japanese Society for Quality Control 14*(2), 147 – 156.

Kauffman, R. J. and L. Diamond (1990, jan). The business value effects of cognitive biases in trading workstation window design. In *System Sciences (HICCS), 1990 23th Hawaii International Conference on*, Volume 4, pp. 469 – 478. IEEE.

Kelly, D. and J. Teevan (2003). Implicit feedback for inferring user preference: A bibliography. *ACM SIGIR Forum 32*(2), 18 – 28.

Kiesler, S. and L. S. Sproull (1986). Response effects in the electronic survey. *Public Opinion Quarterly 50*(3), 402 – 413.

Kirchler, E. and B. Maciejovsky (2002, July). Simultaneous over- and underconfidence: Evidence from experimental asset markets. *Journal of Risk and Uncertainty 25*(1), 65 – 85.

Kleinmuntz, D. N. and D. A. Schkade (1993). Information displays and decision processes. *Psychological Science 4*(4), 221 – 227.

Kosinski, R. J. (2013, September). A literature review on reaction time. Available at http://biae.clemson.edu/bpc/bp/Lab/110/reaction.htm.

Kranz, T. T., F. Teschner, P. Roüast, and C. Weinhardt (2014, February). Identifying individual party preferences in political stock markets. In P. Kommers and P. Isaías (Eds.), *Proceedings of the IADIS International Conference on E-Society*, Volume 12, Madrid, Spain, pp. 162 – 169. IADIS.

Kranz, T. T., F. Teschner, and C. Weinhardt (2014a, June). Combining prediction markets and surveys: An experimental study. In *ECIS 2014 Proceedings*.

Kranz, T. T., F. Teschner, and C. Weinhardt (2014b, January). User heterogeneity in trading systems: Assessing trader's market predisposition via personality questionnaires. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, Volume 1, pp. 1230 – 1239. IEEE Computer Society.

Kranz, T. T., F. Teschner, and C. Weinhardt (2014c, October). Web vs. mobile – comparing trading performance in stationary and mobile settings. *International Journal of E-Services and Mobile Applications 6*(4), 28 – 42.

Krumm, J., N. Davies, and C. Narayanaswami (2008). User-generated content. *Pervasive Computing 7*(4), 10 – 11.

Lakonishok, J. and S. Smidt (1986, September). Volume for winners and losers: Taxation and other motives for stock trading. *The Journal of Finance 41*(4), 951 – 974.

Ledyard, J., R. Hanson, and T. Ishikida (2009). An experimental test of combinatorial information markets. *Journal of Economic Behavior & Organization 69*(2), 182 – 189.

Levin, J. and P. Milgrom (2010, May). Online advertising: Heterogeneity and conflation in market design. *The American Economic Review 100*(2), 603 – 607.

LimeSurvey (2014). Limesurvey. Available at `http://www.limesurvey.org/`.

Lo, A. W., D. V. Repin, and B. N. Steenbarger (2005, March). Fear and greed in financial markets: A clinical study of day-traders. *American Economic Review 95*(2), 352 – 359.

Loewenstein, G. F., E. U. Weber, C. K. Hsee, and N. Welch (2001, March). Risk as feelings. *Psychological Bulletin 127*(2), 267 – 286.

Lucking-Reiley, D. and D. F. Spulber (2001). Business-to-business electronic commerce. *The Journal of Economic Perspectives 15*(1), 55 – 68.

Luckner, S. (2006). Prediction markets: How do incentive schemes affect prediction accuracy? In N. Jennings, G. Kersten, A. Ockenfels, and C. Weinhardt (Eds.), *Negotiation and Market Engineering*, Dagstuhl Seminar Proceedings 06461, Schloss Dagstuhl, Germany.

Luckner, S. (2008). Prediction markets: Fundamentals, key design elements, and applications. In F. Hampe, J. Gricar, A. Pucihar, and G. Lenart (Eds.), *eCollaboration: Overcoming Boundaries through Multi-Channel Interaction*, pp. 236 – 247.

Luckner, S., F. Kratzer, and C. Weinhardt (2005). Stoccer – a forecasting market for the fifa world cup 2006. In *Proceedings of the 4th Workshop on e-Business (WEB 2005)*, pp. 399 – 405.

Luckner, S. and C. Weinhardt (2008, July). Arbitrage opportunities and market-making traders in prediction markets. In *Joint Conference on E-Commerce Technology (CEC'08) and Enterprise Computing, E-Commerce and EServices (EEE '08)*.

Machina, M. J. (1982). "expected utility" analysis without the independence axiom. *Econometrica 50*(2), 277 – 323.

MacLeod, H. (2004, October). *How to be Creative*. ChangeThis.com.

Malhotra, N. K. (1982, March). Information load and consumer decision making. *Journal of Consumer Research 8*(4), 419 – 430.

Malhotra, N. K. (1984, March). Reflections on the information overload paradigm in consumer decision making. *Journal of Consumer Research 10*(4), 436 – 440.

McHugh, P. and A. L. Jackson (2012). Prediction market accuracy: The impact of size, incentives, context and interpretation. *The Journal of Prediction Markets 6*(2), 22 – 46.

McMillan, J. (1994). Selling spectrum rights. *The Journal of Economic Perspectives 8*(3), 145 – 162.

McNees, S. K. (1992, July). How large are economic forecast errors? *New England Economic Review July/August*, 25 – 42.

Milgrom, P. (2011, April). Critical issues in the practice of market design. *Economic Inquiry 49*(2), 311 – 320.

Muntermann, J. and L. Janssen (2005). Assessing customers' value of mobile financial information services: Empirical-based measures. In *ICIS 2005 Proceedings*, pp. 617 – 628.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika 78*(3), 691 – 692.

Neumann, D. G. (2007). *Market Engineering – A Structured Design Process for Electronic Markets*. Dissertation, Universität Karlsruhe (TH) - IISM, Englerstr. 14, 76131 Karlsruhe.

O'Connor, P. and F. Zhou (2008). The tradesports nfl prediction market: An analysis of market efficiency, transaction costs, and bettor preferences. *The Journal of Prediction Markets 2*(1), 45 – 71.

Odean, T. (1998, October). Are investors reluctant to realize their losses? *The Journal of Finance 53*(5), 1775 – 1798.

O'Hara, M. (1995). *Market Microstructure Theory*. Cambridge, MA: Blackwell Publishers.

Oliven, K. and T. A. Rietz (2004, March). Suckers are born but markets are made: Individual rationality, arbitrage, and market efficiency on an electronic futures market. *Management Science 50*(3), 336 – 351.

O'Rand, A. M. and M. L. Krecker (1990). Concepts of the lyfe cycle: Their history, meanings, and uses in the social sciences. *Annual Review of Sociology 16*, 241 – 262.

Osborn, A. F. (1953). *Applied Imagination*. Scribner's.

Quinn, R. E. and K. Cameron (1983, January). Organizational life cycles and shifting criteria of effectiveness: Some preliminary evidence. *Management Science 29*(1), 33 – 51.

Ranchhod, A. and F. Zhou (2001). Comparing respondents of e-mail and mail surveys: Understanding the implications of technology. *Marketing Intelligence & Planning 19*(4), 254 – 262.

Riordan, R. and A. Storkenmaier (2012, June). Latency, liquidity and price discovery. *Journal of Financial Markets 15*(4), 416 – 437.

Roth, A. E. (2002, July). The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica 70*(4), 1341 – 1378.

Roth, A. E. (2008, March). What have we learned from market design? *The Economic Journal 118*(527), 285 – 310.

Roth, A. E. (2010, September). What have we learned from market design? *El Trimestre Economico*.

Rothschild, D. (2009). Forecasting elections: Comparing prediction markets, polls, and their biases. *Public Opinion Quarterly 73*(5), 895 – 916.

Sakamoto, Y., M. Ishiguro, and K. G. (1986). *Akaike Information Criterion Statistics*. Reidel Publishing Company.

Schkade, D. A. and D. N. Kleinmuntz (1994). Information displays and choice processes: Differential effects of organization, form, and sequence. *Organizational Behavior and Human Decision Processes 57*(3), 319 – 337.

Schmiedl, G., M. Seidl, and K. Temper (2009). Mobile phone web browsing – a study on usage and usability of the mobile web. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '09, New York, NY, USA, pp. 1 – 2. ACM.

Schnizler, B. (2008). Mace: A multi-attribute combinatorial exchange. In H. Gimpel, N. R. Jennings, G. E. Kersten, A. Ockenfels, and C. Weinhardt (Eds.), *Negotiation, Auctions, and Market Engineering*, Volume 2 of *Lecture Notes in Business Information Processing*, pp. 84 – 100. Springer Berlin Heidelberg.

Schnizler, B., D. Neumann, D. Veit, and C. Weinhardt (2008). Trading grid services – a multi-attribute combinatorial approach. *European Journal of Operational Research 187*(3), 943 – 961.

Schnizler, B., D. Neumann, and C. Weinhardt (2004). Resource allocation in computational grids – a market engineering approach. In *Proceedings of the Third Workshop on e-Business (WEB 2004)*, Washington D. C., USA.

Schönfeld, A. and C. Block (2010). A meta-framework for agile development of soa market platforms. *SSRN Working Paper* (1600078).

Segerstrom, P. S., T. C. A. Anant, and E. Dinopoulos (1990, December). A schumpeterian model of the product life cycle. *The American Economic Review 80*(5), 1077 – 1091.

Selhorst, S. (2006). Prioritizing software requirements with kano analysis. *The Pragmatic Marketer Magazine 4*(3), 24 – 29.

Seru, A., T. Shumway, and N. Stoffman (2010). Learning by trading. *The Review of Financial Studies 23*(2), 705 – 739.

Servan-Schreiber, E., J. Wolfers, D. M. Pennock, and B. Galebach (2004, September). Prediction markets: Does money matter? *Electronic Markets 14*(3), 243 – 251.

Seuken, S., K. Jain, and D. C. Parkes (2010). Hidden market design. In *AAMAS '10 Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, Volume 1, pp. 1661–1662.

Shapira, Z. and I. Venezia (2001, August). Patterns of behavior of professionally managed and independent investors. *Journal of Banking & Finance 25*(8), 1573 – 1587.

Shefrin, H. and M. Statman (1985, Dezember). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance 40*(3), 777 – 790.

Sjöberg, L. (2009). Are all crowds equally wise? a comparison of political election forecasts by experts and the public. *Journal of Forecasting 28*(1), 1 – 18.

Slamka, C., A. Soukhoroukova, and M. Spann (2008). Event studies in real- and play-money prediction markets. *The Journal of Prediction Markets 2*(2), 53 – 70.

Smith, V. L. (1982, December). Microeconomic systems as an experimental science. *The American Economic Review 72*(5), 923 – 955.

Snowberg, E. and J. Wolfers (2010). Explaining the favorite-longshot bias: Is it risk-love or misperceptions? *NBER Working Paper 15923*, 1 – 29.

Snowberg, E., J. Wolfers, and E. Zitzewitz (2007, May). Partisan impacts on the economy: Evidence from prediction markets and close elections. *The Quarterly Journal of Economics 122*(2), 807 – 829.

Spann, M. and B. Skiera (2003, October). Internet-based virtual stock markets for business forecasting. *Management Science 49*(10), 1310 – 1326.

Speier, C. and M. G. Morris (2003, Sep). The influence of query interface design on decision-making performance. *MIS Quarterly 27*(3), 397 – 423.

Stanovich, K. E. and R. F. West (2000, October). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences 23*(5), 645 – 726.

Stark, J. (2011). *Product Lifecycle Management*, pp. 1 – 16. Springer London.

Stathel, S., S. Luckner, F. Teschner, C. Weinhardt, A. Reeson, and S. Whitten (2009). *AKX – An Exchange for Predicting Water Dam Levels in Australia*, pp. 78–90. Springer Berlin Heidelberg.

Stathel, S., F. Teschner, T. Kullnig, T. Kranz, C. van Dinther, and C. Weinhardt (2010). Innovation assessment via enterprise information markets. In *Proceedings of the 1st International Conference on IT-enabled Innovation in Enterprise*, pp. 206 – 218.

Stathel, S., C. van Dinther, and A. Schönfeld (2009, February). Service innovation with information markets. In *WIRTSCHAFTSINFORMATIK PROCEEDINGS 2009*, Volume 1, Vienna, Austria, pp. 825 – 843.

Straub, T., H. Gimpel, and F. Teschner (2014, January). The negative effect of feedback on performance in crowd labor tournaments. In J. Nickerson and T. Malone (Eds.), *Collective Intelligence 2014: Proceedings*.

Straub, T., H. Gimpel, F. Teschner, and C. Weinhardt (2014, June). Feedback and performance in crowd work: A real effort experiment. In *ECIS 2014 Proceedings*.

Streufert, S., M. J. Driver, and K. W. Haun (1967). Components of response rate in complex decision-making. *Journal of Experimental Social Psychology 3*(3), 286 – 295.

Stroustrup, B. (1990). Quote. Available at `http://www.stroustrup.com/bs_faq.html#really-say-that`.

Subrahmanyam, A. (1991). Risk aversion, market liquidity, and price efficiency. *The Review of Financial Studies 4*(3), 417 – 441.

Sunstein, C. R. (2006). *Information Markets: A New Way of Making Decisions*, Chapter Deliberation and Information Markets, pp. 67 – 100. Washington, D. C.: AEI-Brookings Press.

Surowiecki, J. (2004). *The Wisdom of Crowds*. Random House LLC.

Tan, C.-H., H.-H. Teo, and I. Benbasat (2010, June). Assessing screening and evaluation decision support systems: A resource-matching approach. *Information Systems Research 21*(2), 305 – 326.

Teo, T. S. H. and C. Ranganathan (2004). Adopters and non-adopters of business-to-business electronic commerce in singapore. *Information & Management 42*(1), 89 – 102.

Teschner, F. (2012, July). *Forecasting Economic Indices — Design, Performance, and Learning in Prediction Markets*. Dissertation, Fakultät für Wirtschaftswissenschaften am Karlsruher Institut für Technologie (KIT).

Teschner, F., T. T. Kranz, and C. Weinhardt (2012, March). Decision behavior and performance in mobile trading applications. In A. Back, M. Bick, M. Breunig, K. Pousttchi, and F. Thiesse (Eds.), *MMS 2012: Mobile und Ubiquitäre Informationssysteme*, Volume P-202, Bonn, pp. 122 – 127. Gesellschaft für Informatik (GI), Bonn: Köllen Druck+Verlag GmbH, Bonn.

Teschner, F., T. T. Kranz, and C. Weinhardt (2014). The impact of customizable market interfaces on trading performance. *Electronic Markets*.

Teschner, F., A. Mazarakis, R. Riordan, and C. Weinhardt (2011, December). Participation, feedback & incentives in a competitive forecasting community. In *ICIS 2011 Proceedings*.

Teschner, F., R. Riordan, and C. Weinhardt (2011). Behavioral ict — risk, cognition and information. In T. Eymann (Ed.), *Proceedings of the Doctoral Consortium of the WI 2011*, pp. 174 – 183.

Teschner, F., S. Stathel, and C. Weinhardt (2011). A prediction market for macro-economic variables. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on*, pp. 1 – 9.

Teschner, F., F. Wagenschwanz, and C. Weinhardt (2012). Analysis of the disposition effect: Asymetry and prediction accuracy. *The Journal of Prediction Markets 7*(1), 27 – 42.

Teschner, F. and C. Weinhardt (2012). Evaluating hidden market design. In P. Coles, S. Das, S. Lahaie, and B. Szymanski (Eds.), *Auctions, Market Mechanisms, and Their Applications*, Volume 80 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 5 – 17. Springer Berlin Heidelberg.

Teschner, F. and C. Weinhardt (2014, July). A macroeconomic forecasting market. *Journal of Business Economics*, 1 – 25. http://dx.doi.org/10.1007/s11573-014-0741-5.

Tetlock, P. E. (2005). *Expert political judgment: How good is it? How can we know?* Princeton University Press.

Teubner, T. (2013, Dezember). *Social Preferences under Risk: Peer Types and Relationships in Economic Decision Making*. Dissertation, KIT-IISM.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science 4*(3), 199 – 214.

Thaler, R. H. (Ed.) (1993). *Advances in Behavioral Finance*, Volume 1. Russel Sage Foundation.

Thaler, R. H. and H. M. Shefrin (1981, April). An economic theory of self-control. *Journal of Political Economy 89*(2), 392 – 406.

Thaler, R. H., C. R. Sunstein, and J. P. Balz (2010). Choice architecture. *SSRN Working Paper 1583509*, 1 – 17.

Thaler, R. H. and W. T. Ziemba (1988). Anomalies: Parimutuel betting markets: Racetracks and lotteries. *The Journal of Economic Perspectives 2*(2), 161 – 174.

Tversky, A. and D. Kahneman (1974, September). Judgment under uncertainty: Heuristics and biases. *Science 185*(4157), 1124 – 1131.

Uhlaner, C. J. and B. Grofman (1986). The race may be close but my horse is going to win: Wish fulfillment in the 1980 presidential election. *Political Behavior 8*(2), 101 – 129.

van Dijck, J. (2009, October). Users like you? theorizing agency in user-generated content. *Media, Culture & Society 31*(1), 41 – 58.

van Witteloostuijn, A. and K. Muehlfeld (2008). Trader personality and trading performance – a framework and financial market experiment. *Discussion Paper Series / Tjalling C. Koopmans Research Institute 8*(24), 1 – 44.

Varian, H. R. (2002, August). Avoiding the pitfalls when economics shifts from science to engineering. Technical Report Aug 29, 2002, New York Times, New York, NY, USA.

Venkatesh, V. (2000, December). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information Systems Research 11*(4), 342 – 365.

Venkatesh, V. and F. D. Davis (2000, February). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science 46*(2), 186 – 204.

Vernon, R. (1966). International investment and international trade in the product cycle. *Quarterly Journal of Economics 80*(2), 190 – 207.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences 22*(2), 219–240.

Vessey, I. (1994). The effect of information presentation on decision making: A cost-benefit analysis. *Information & Management 27*(2), 103 – 119.

Vessey, I. and D. Galletta (1991). Cognitive fit: An empirical study of information acquisition. *Information Systems Research 2*(1), 63 – 84.

von Hayek, F. A. (1945, September). The use of knowledge in society. *The American Economic Review 35*(4), 519 – 530.

Wärneryd, K.-E. (1996, December). Risk attitudes and risky behavior. *Journal of Economic Psychology 17*(6), 749 – 770.

Weber, M. and C. F. Camerer (1998, January). The disposition effect in securities trading: an experimental analysis. *Journal of Economic Behavior & Organization 33*(2), 167 – 184.

Weber, M. and F. Welfens (2007). An individual level analysis of the disposition effect: Empirical and experimental evidence. Technical Report 07-45, DFG SFB 504, University of Mannheim.

Weinhardt, C. and H. Gimpel (2007). Market engineering: An interdisciplinary research challenge. In N. Jennings, G. Kersten, A. Ockenfels, and C. Weinhardt (Eds.), *Negotiation and Market Engineering*, Number 06461, Dagstuhl, Germany, pp. 1 – 15. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.

Weinhardt, C., C. Holtmann, and D. Neumann (2003). Market-engineering. *Wirtschaftsinformatik 45*(1), 635 – 640.

Wells, W. D. and G. Gubar (1966, November). Life cycle concept in marketing research. *Journal of Marketing Research 3*(4), 355 – 363.

Wilson, R. (2002, July). Architecture of power markets. *Econometrica 70*(4), 1299 – 1340.

Wolfers, J. and E. Zitzewitz (2004, May). Prediction markets. *Journal of Economic Perspectives 18*(2), 107 – 126. http://www.nber.org/papers/w10504.

Wolfers, J. and E. Zitzewitz (2006, March). Prediction markets in theory and practice. Working Paper 12083, National Bureau of Economic Research. http://www.nber.org/papers/w12083.

Yang, S., Y. Hsu, and C. Tu (2012, September). How do traders influence investor confidence and trading volume? a dyad study in the futures market. *Emerging Markets Finance & Trade 48*(Supplement 3), 23 – 34.

Zhang, S., M. Wagener, A. Storkenmaier, and C. Weinhardt (2011). The quality of electronic markets. In *Proceedings of the 44th Hawaii International Conference on System Sciences*, pp. 1530 – 1605.

Zhang, Y. (1999). Using the internet for survey research: A case study. *Journal of the American Society for Information Science 51*(1), 57 – 68.

Zitzewitz, E. (2006, July). Price discovery among the punters: Using new financial betting markets to predict intraday volatility. *SSRN Working Paper 925742*.

List of Abbreviations

AfD	**Al**ternative **f**ür **D**eutschland	55
AIC	**A**kaike **I**nformation **C**riterion	76
ANOVA	**An**alysis **o**f **Va**riance	123
ARIMA	**A**uto**r**egressive **I**ntegrated **M**oving **A**verage	29
CDA	**C**ontinuous **D**ouble **A**uction	31
CDU	**C**hristlich **D**emokratische **U**nion	54
CRA	**C**ognitive **R**eflection **A**bilities	6
CRT	**C**ognitive **R**eflection **T**est	66
CSU	**C**hristlich **S**oziale **U**nion	54
DAX	**D**eutscher **A**ktieninde**x**	47
DE	**D**isposition **E**ffect	113
EIX	**E**conomic **I**ndicator e**X**change	52
EMA	**E**IX **M**arket **A**pp	53
EMH	**E**fficient-**M**arket **H**ypothesis	30
ERS	**E**motion **R**egulation **S**trategies	6
ERQ	**E**motion **R**egulation **Q**uestionnaire	67
FDP	**F**reiheitlich **D**emokratische **P**artei	54
GDP	**G**ross **D**omestic **P**roduct	130
GMT	**G**reenwich **M**ean **T**ime	48
Grüne	Bündnis ’90/Die **Grüne**n	54
HCI	**H**uman-**C**omputer **I**nteraction	96
IPSM	**I**owa **P**residential **S**tock **M**arket	84
KAPP	**K**urspiloten **App**lication	50
MD	Method of **M**ajority **D**ecision	54
mfx	**M**arginal **E**ffect	75
NBA	**N**ational **B**asketball **A**ssociation	114
NFL	**N**ational **F**ootball **L**eague	114
PGR	**P**roportion of **G**ains **R**ealized	120
Piraten	**Piraten**partei	54
PIX	**P**olitical **I**ndicator e**X**change	53
PIX€	**P**olitical **I**ndicator e**X**change €	46
PLR	**P**roportion of **L**osses **R**ealized	120
PR	**P**roportional **R**epresentation	54
P€	**P**lay-money €	46
PSM	**P**olitical **S**tock **M**arket	84
RA	**R**isk **A**version	6

List of Figures

List of Tables