

Judgmental Biases in the Evaluation of Innovations: Experiments with Information Markets

Vom Promotionsausschuss der
Technischen Universität Hamburg-Harburg
zur Erlangung des akademischen Grades
Doktor der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol)
genehmigte Dissertation

von
Jan-Paul Lüdtke

aus
Bremen

2015

1. Gutachter: Professor Dr. Christian Lüthje
2. Gutachter: Professor Dr. Nikolaus Franke

Tag der mündlichen Prüfung: 13.5.2015

Contents

List of Figures	V
List of Tables	VIII
List of Abbreviations	X
I. Introduction	1
1. Introduction	2
1.1. Research scope: uncertainty and bias in the evaluation of innovations . . .	2
1.2. Research objectives	5
1.3. Research structure and organization	6
II. Conceptual Background	9
2. The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty	10
2.1. Relevant characteristics of innovations	10
2.2. Uncertainty in innovation	13
2.2.1. The concept of uncertainty in innovation	13
2.2.2. Drivers of uncertainty in innovation	15
2.2.3. The impact of uncertainty on innovation-related decision making .	20
2.3. Reducing uncertainty	23
2.3.1. Improving the information base	25
2.3.2. Filtering valid information	33
3. Information Markets for Innovation Evaluation	48
3.1. Foundations of information markets	48
3.2. Applications of information markets for innovation evaluation	54
3.3. Designing information markets for innovation evaluation	61
3.3.1. Prediction objects	61
3.3.2. Participants	64
3.3.3. Trading mechanisms	66
3.3.4. Incentives	69

4. Judgmental Biases in Innovation Evaluation	72
4.1. (Ir)rationality in economic behavior	72
4.2. Fundamentals of bias in decision making	74
4.2.1. Biases from heuristics and cognitive processing	75
4.2.2. Biases from motivation	76
4.3. Systematic literature review of research on biases in innovation management	78
4.3.1. Methodology	79
4.3.2. Results	80
4.4. Judgmental biases in information markets for innovation evaluation . . .	91
4.5. Summary	93
 III. Empirical Studies	 95
5. Research Framework	96
5.1. Focus on overconfidence	96
5.2. Level of analysis	98
5.3. Methodology	99
5.4. Research process	103
5.4.1. Individual confidence manipulation	104
5.4.2. Overconfidence and individual behavior in information markets .	105
5.4.3. Overconfidence and the prediction quality of information markets	107
 6. Treatment Check	 109
6.1. Experimental design and implementation	109
6.2. Innovation evaluation tasks	111
6.3. Results	112
 7. Experiment 1: Overconfidence and Individual Behavior in Information Mar-	 115
kets	
7.1. Development of hypotheses	115
7.2. Experimental design	120
7.2.1. Participants	120
7.2.2. Implementation	121
7.2.3. Instructions	126
7.3. Sample and data preparation	126
7.3.1. Missing data	126
7.3.2. Variable operationalization and construct validity	128
7.3.3. Data adequacy and statistical model	131
7.4. Results	134
7.5. Discussion	141
 8. Experiment 2: Overconfidence and the Prediction Quality of Information	 144
Markets	

Contents

8.1. Development of hypotheses	145
8.1.1. Overconfidence and prediction quality of information markets . .	145
8.2. Experimental design	151
8.2.1. Participants	151
8.2.2. Innovation evaluation tasks	151
8.2.3. Implementation	153
8.2.4. Instructions	157
8.3. Data preparation	157
8.3.1. Variable operationalization and construct validity	157
8.3.2. Data adequacy and statistical model	159
8.4. Results	160
8.4.1. Overconfidence and the prediction quality of information markets	161
8.4.2. An additional setting where information for treated individuals was provided at a cost	172
8.5. Discussion	178
IV. Synthesis	183
9. Synthesis	184
9.1. Summary of results	184
9.2. Limitations	186
9.3. Implications for innovation management	189
9.4. Research outlook	195
V. Bibliography	198
Bibliography	199
VI. Appendix	223
VII. Curriculum Vitae	250

List of Figures

1.1. Structure and process of the dissertation (Source: Own depiction)	7
2.1. The fuzzy front-end of the innovation process (Source: Own depiction based on Herstatt et al. (2004))	16
2.2. Reducing uncertainty and increasing the validity of innovation evalua- tion by improving the information base and applying effective evaluation (Source: Own depiction)	24
3.1. Information market process example, from event definition to stock pay- out (Source: Own depiction)	50
3.2. Prediction quality in global political information markets between 1988 and 2000 (Source: Berg et al. (2008))	52
3.3. Previous application of information markets for innovation evaluation . .	60
4.1. Classical model of motivation from motivational psychology (Source: Rhein- berg (1997))	77
4.2. Number of articles on biases in innovation management, by bias and year (Source: Own depiction)	81
5.1. Research process and level of analysis of the three empirical studies (Source: Own depiction)	104
6.1. Experimental procedure, by treatment group, in the treatment check (Source: Own depiction)	110
6.2. Change in (over-)confidence before and after treatment in the treatment check (Source: Own depiction)	113
7.1. Laboratory set-up for the first experiment (Source: Own depiction) . . .	122
7.2. Experimental procedure, by treatment group, in the first experiment (Source: Own depiction)	123

List of Figures

7.3. Information market interaction screen for subjects under LC-treatment (Source: Own depiction)	125
7.4. Distributions and residual normality checks for the Hierarchical Linear Model (HLM) analysis	133
7.5. Development of market transactions, by treatment condition, over the market's running time (Source: Own depiction)	139
8.1. Design for the second experiment, by experimental setting: Number of cases by treatment and experimental setting	153
8.2. Information market interaction interface for subjects in the overconfidence (top) and low confidence treatment (bottom) groups, in the free informa- tion condition	156
8.3. Interaction effect of noise trader pre-market error and the presence of overconfident traders on market error (Source: Own depiction)	164
8.4. Indicators of trading behavior by treated subjects and treatment condition (Source: Own depiction)	165
8.5. Indicators of trading behavior by noise traders and treatment condition (Source: Own depiction)	168
8.6. Interaction effect of treated subjects' pre-market errors and treatment conditions on their post-market errors	171
9.1. Articles for literature analysis	224
9.2. Articles for literature analysis continued	225
9.3. Articles for literature analysis continued	226
9.4. Articles for literature analysis continued	227
9.5. Treatment Check - Evaluation questions	228
9.6. Treatment Check - Evaluation questions continued	229
9.7. Experiment 1 - Instructions page 1	230
9.8. Experiment 1 - Instructions page 2	231
9.9. Experiment 1 - Instructions page 3	232
9.10. Experiment 1 - Pre-experimental questionnaire	233
9.11. Experiment 1 - Evaluation questions and information-market questions .	234
9.12. Experiment 1 - Post-experimental questionnaire page 1	235
9.13. Experiment 1 - Post-experimental questionnaire page 2	236
9.14. Experiment 2 - Instructions page 1	237
9.15. Experiment 2 - Instructions page 2	238

List of Figures

9.16. Experiment 2 - Instructions page 3	239
9.17. Experiment 2 - Instructions page 4	240
9.18. Experiment 2 - Instructions page 5	241
9.19. Experiment 2 - Instructions page 6	242
9.20. Experiment 2 - Pre-experimental questionnaire	243
9.21. Experiment 2 - Evaluation questions and information-market questions .	244
9.22. Experiment 2 - Evaluation questions and information-market questions continued	245
9.23. Experiment 2 - Data adequacy check for the basic experimental setting .	246
9.24. Experiment 2 - Data adequacy check for the additional experimental setting	247
9.25. Experiment 2 - Post-experimental questionnaire page 1	248
9.26. Experiment 2 - Post-experimental questionnaire page 2	249

List of Tables

2.1. Accessibility and goal-compatibility fit for FTF, Delphi and NGT, and information markets (Source: Own depiction)	42
4.1. Journals selected for the systematic literature review	80
6.1. Treatment check results	112
7.1. Number of subjects per recorded market periods yield	127
7.2. Operationalization of variables included in the statistical analysis for Experiment 1	129
7.3. Measurement of constructs for covariates	130
7.4. Correlation, means and standard deviations for variables at Level 1 . . .	131
7.5. Correlation, means and standard deviations for variables at Level 2 . . .	132
7.6. The impact of confidence on the time of first trade (in s)	135
7.7. The impact of confidence on total stock turnover	137
7.8. The impact of confidence on net stock turnover towards initial estimate .	138
7.9. The impact of confidence on the delta between pre- and post-market evaluation	140
7.10. Average outcomes for dependent variable by treatment condition	141
8.1. Operationalization of variables included in the statistical analysis for Experiment 2	158
8.2. Measurement of constructs for control variables	159
8.3. Impact of the presence of overconfident individuals on market prediction error in the basic experimental setting	162
8.4. The impact of treatment condition on treated subjects' post-market error and post-market prediction improvements by noise traders in the basic experimental setting	170

List of Tables

8.5. Regression results regarding individual post-market prediction errors by treated subjects and individual prediction improvements by noise traders	173
8.6. Impact of the presence of overconfident individuals on market prediction error when information comes at a cost	175
8.7. The impact of treatment condition on post-market error by treated subjects and post-market prediction improvements by noise traders when information comes at a cost	177
8.8. Individual and market prediction errors in the basic experimental setting	179
8.9. Individual and market prediction error in the additional experimental setting	181

List of Abbreviations

AATD	Artificial-Agent Trade Direction
AMJ	Academy of Management Journal
CJF	Combined Judgmental Forecast
FPRA	Financial-Product Risk-Attitude
FTF	Face-to-Face Meeting
HLM	Hierarchical Linear Model
HSX	Hollywood Stock Exchange
IGM	Interactive Group Method
JBV	Journal of Business Venturing
JM	Journal of Marketing
JMR	Journal of Marketing Research
JPIM	Journal of Product Innovation Management
LA	Loss Aversion
LC	Low Confidence
M	Million
ManSci	Management Science
MANOVA	Mean Analysis Of Variance
MAR	Missing At Random
MCAR	Missing Completely At Random
MNAR	Missing Not At Random
NGT	Nominal-Group Technique
OC	Overconfidence
OrgSci	Organization Science
PDI	Product Domain Involvement
PT	Prospect Theory
RP	Research Policy
RSE	Root Square Error
R&D	Research and Development

List of Tables

SMJ	Strategic Management Journal
STOC	Securities Trading Of Concepts
TD	Trading Distance

Part I.

Introduction

1. Introduction

1.1. Research scope: uncertainty and bias in the evaluation of innovations

Successful innovation is a vital activity for any organization. Continuous innovation is a crucial factor in a company's long-term success because it allows organizational output to adapt to evolving demands. However, two fundamental observations in innovation management have been that (1) the development of innovation comes with a considerable risk of failure and (2) it is no trivial task to make a truthful assessment of these risks (Mullins and Sutherland 1998; Schmidt et al. 2009).

A number of researchers have explored why it seems particularly difficult to evaluate the success potential of innovations (Reid and De Brentani 2004; Ozer 2005; Leiponen and Helfat 2010). If there were a common denominator to these research findings, it would be that any innovative undertaking must embrace uncertainty (Jalonen 2011). For example, when deciding to start development on an innovation, it often remains uncertain whether

1. ...the innovator has truly grasped the problem that he/she aims to solve. It may be uncertain which customers to listen to (Christensen and Bower 1996) and how to properly understand the responses of (potential) customers (Billeter et al. 2011)
2. ...the innovator's perception of his/her problem solving capability matches reality. For example, it may be uncertain whether the innovator has at his/her command the resources required to develop the targeted solution (Leifer et al. 2001).
3. ...the problem will persist even after the innovator provides a solution. Especially in rapidly changing market environments, it may be uncertain whether the needs identified at the outset of an innovation project will match customer requirements once the solution is ready. Such uncertainty is particularly relevant today because

1. Introduction

(1) innovation development times are getting longer due to increasingly complex technologies and (2) market environments and customer needs are evolving more rapidly due to more liberal markets (Griffin 1997).

To properly evaluate any innovation requires sufficiently good estimates for the probability distributions of expected returns and investments; however, the above-mentioned uncertainties, among others, make the estimates of those figures particularly prone to errors. Looking back at history, a plethora of examples is available to demonstrate how innovators have often failed to correctly estimate costs or returns (Stevens and Burley 2003). Even the biographies of individuals who have at times been hailed as the most successful corporate innovators are equally marked by failures that demonstrate quite the opposite (Denrell and Fang 2010).

Hence, one of the core tasks of innovation management is to decrease uncertainty regarding the potential of innovative endeavors so that innovators and investors can make better decisions in starting and executing these projects. Scholars have devoted much attention to identifying, developing, and testing methods to increase the likelihood of returns and drain less resources in the course of innovation development (Evanschitzky et al. 2012). Schmidt et al. (2009) found that proficiency in evaluating an innovation has a significantly positive impact on the potential of new products. Sound information, evaluation and decisions are particularly important during the early phases of an innovation project because early choices have a considerably larger impact on an innovation's success potential than later ones (Stockstrom and Herstatt 2008). Consider, for example, an innovation project in which the innovator makes an erroneous assumption at the outset of development. If the innovator or investor fails to notice the error, the innovator may end up with a failed product and the investor will be saddled with a gloomy return on investment, even if all subsequent actions are perfectly executed.

Several strands of research have introduced the notion that the validity of an innovation's evaluation improves when decision makers access and utilize heterogeneous information residing in multiple domains within and outside the innovating organization (Gassmann 2006a; Gupta et al. 2007; Poetz and Schreier 2012). A relatively new and promising stream of research, for example, has explored novel methods to harness the **wisdom of crowds** to predict the outcomes of future events like innovations (Surowiecki 2005).

A central driver behind the rising recognition and fame of crowd wisdom lies in an unprecedented increase in virtual interaction and communication involving large parts

1. Introduction

of the global population without significant geographical or social boundaries. This is done via electronic networks. Like never before, organizations can implement large-scale aggregations of heterogeneous information, which may even originate outside the organizations' boundaries, with the goal of reducing innovation-related uncertainty. Furthermore, many companies have actively opened up their innovation processes to external informants (Dahlander and Gann 2010). For example, companies have started to source problem-specific advice for difficult technological challenges from highly specialized experts from all over the world via open innovation competitions or knowledge exchanges (Boudreau et al. 2011). Empirical results from lab experiments and field applications indicate that tapping heterogeneous individual information and expectations yields superior evaluation and forecasting results over traditional methods of decision support. Mechanisms that draw from the **wisdom of crowds** have frequently outperformed established methods like extrapolation or expert judgment for forecasting the success potential of new products (Chen and Plott 2002; Spann and Skiera 2003a).

In the specific context of innovation evaluation, **information markets** have drawn particular attention as a promising tool for predicting the success of innovative ideas and concepts (Spears et al. 2009; Soukhoroukova et al. 2012) or new products (Dahan et al. 2010). Information markets are interactive market platforms that incentivize participants depending on their ability to forecast the outcome of uncertain events. These markets have been successfully applied in highly innovative companies such as Google or Microsoft by tapping the wisdom of crowds (Surowiecki 2005; Cowgill et al. 2008). The success of information markets is underpinned by the pooling heterogeneous beliefs that contribute relevant information regarding the prediction task. Drawing on rational expectations theory, information markets align subject incentives and predictive quality. That is, information markets use virtual stocks to represent the future success of innovations or new products. The participants trade shares of these virtual stocks, such that the resulting stock prices indicate the success potential of innovative ideas, concepts, or new products (Soukhoroukova et al. 2012). If individual participants' beliefs are better than current group predictions, they increase the aggregate group predictions along with their expected payouts from trading (Arrow et al. 2008). Lab experiments show that these structured methods of predicting outcomes of uncertain events can outperform unstructured methods such as face-to-face meetings (Graefe and Armstrong 2011).

Ultimately, multiple strands of research from the fields of innovation management, orga-

1. Introduction

nizational science, and entrepreneurship have documented that **predicting the future success of innovations or new products is often prone to assessment biases** if the evaluation originates from human judgments (Hayward et al. 2006; Karniouchina 2011; Peer 2012). Humans often depart from making evaluations that align with organizational success. On the one hand, such departures may be caused by agency problems arising when evaluators pursue different goals than the entity calling for the evaluation. These goals may result from conscious formation or from motivational biases such as the need to maintain high levels of self-esteem. On the other hand, cognitive biases frequently prevent human agents from making valid evaluations. Even when evaluators' intentions for revealing information may be completely aligned with organizational goals, personal information environments or information processing capabilities may hinder them from revealing valid information.

While initial results have shown that information markets often outperform traditional methods, more recent findings question their robustness in the presence of biased participants (Sonnemann 2008; Seybert and Bloomfield 2009). Many researchers have stressed the need to investigate the robustness of information markets against the impact of judgmental biases (e.g. Wu et al. (2008) and Spears et al. (2009)).

1.2. Research objectives

The **main goal** of this thesis is to foster understanding about the impact of biases on the evaluation of innovation via information markets. The two central objectives of the study are as follows. First, a conceptual background must be constructed. From there we can derive a field for relevant empirical investigation.

Constructing the **conceptual background** encompasses three central elements. We first need to arrive at common understanding of why the aggregation of heterogeneous beliefs is particularly important for increasing innovation success. Then, a detailed introduction to the methodological background and current findings on information markets in the context of innovation evaluation is necessary. This will assist in deriving relevant facets for designing and evaluating information markets for subsequent empirical research. Last, a thorough understanding must be achieved about why and how judgmental biases negatively impact decision making in innovation evaluation tasks. Here, we will describe the relevant biases and their important characteristics.

1. Introduction

The **empirical investigation** includes three elements that contribute to the main goal: First, the empirical investigation must be focused on an area that can be effectively addressed in the context of this study. This will be achieved by placing the empirical focus on judgmental bias, which is likely to have the strongest detrimental effect on the quality of information market outcomes in the context of innovation evaluation. Second, we need to develop an understanding of how the selected bias, (*overconfidence*), impacts individual behavior in the context of information markets. Third, a feasible study must be designed and carried out to investigate the impact of individual behaviors on the quality of information market results and to investigate the impact of overconfidence on the quality of innovation evaluation via information markets.

The findings of this study will ultimately provide an understanding about how overconfidence, as a highly relevant judgmental bias in the context of innovation evaluation, will likely impact the outcomes of information markets for use in innovation evaluation. The findings will be empirically supported and connected to existing theoretical and empirical findings in order to provide novel understanding regarding the mechanics of overconfidence in the context of innovation evaluation via information markets. The discussions will provide ideas for future research in domains related to judgmental biases and innovation evaluation. Furthermore, we will provide suggestions for how decision makers can address the impact of overconfidence to increase the quality of information market outcomes, and consequently, the potential for innovation success.

1.3. Research structure and organization

This thesis is divided into four parts, which are visualized in Figure 1.1. The current section concludes the introduction, which represents the first part.

In **Part 2**, we introduce the conceptual foundations upon which this thesis rests. First, relevant characteristics of innovations are introduced and the difficulty of properly evaluating innovation success potential is related to the concept of uncertainty in innovation. We then highlight relevant sources for gathering, and methods for filtering, innovation-related information. The importance of human expertise from different backgrounds and its integration in evaluation processes will be particularly highlighted.

From there, information markets in the context of innovation evaluation will be discussed. We will provide a thorough introduction to the information markets method, highlight previous applications of it in the domain of innovation evaluation, and discuss

1. Introduction

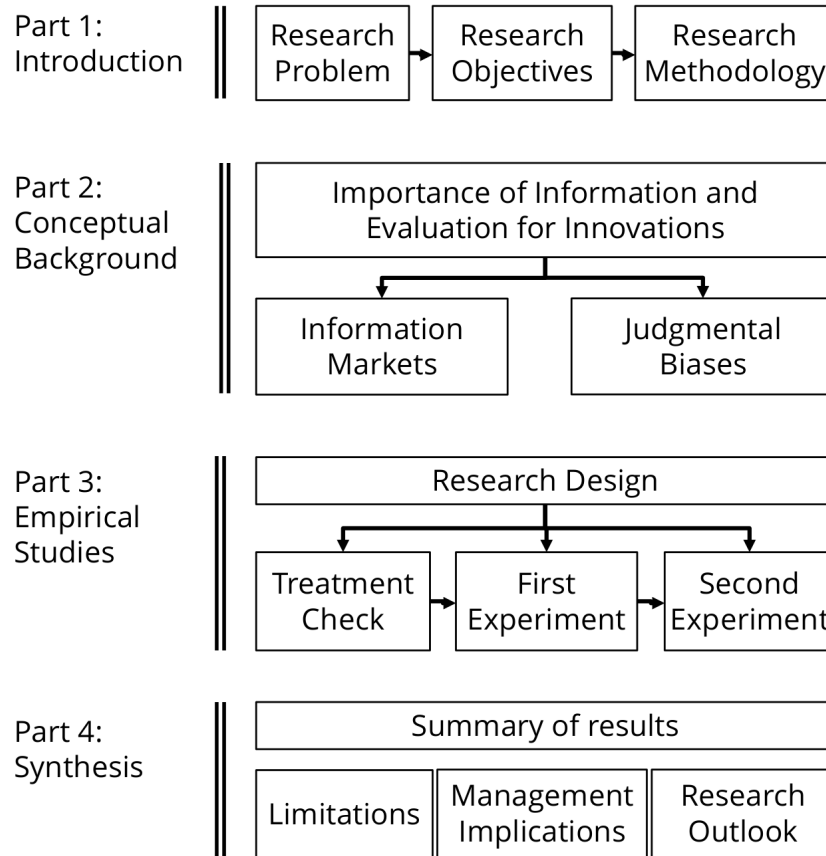


Figure 1.1.: Structure and process of the dissertation (Source: Own depiction)

the design options.

Finally, human difficulties in rendering judgments under uncertainty will be illustrated and explained with a focus on innovation management. A literature review will be conducted to identify judgmental biases, to which difficulties can largely be attributed when making innovation-related decisions. We will explore the biases' origins, their underlying mechanisms and highlight their consequences in innovation-related decisions. We ultimately highlight how judgmental biases may particularly influence the results of information markets in the context of innovation evaluation.

Part 3 covers the empirical work. We begin this section by framing the empirical research. We explain our specific focus on the overconfidence bias and the methodology for studying it in the context of information markets. The empirical work consists of three parts. First, we briefly document the applicability of a treatment to experimentally in-

1. Introduction

duce overconfidence for subsequent use in the information market experiments. Then, in the first experiment, the impact of overconfidence on individual behavior in information markets is studied. This experiment specifically focuses on how overconfidence impacts trading behavior at the subject level. In the second experiment, we study the impact of overconfident individuals' trading behavior on the prediction quality of information markets. Instead of focusing on individual behavior as the dependent variable, the second experiment aims to provide insight into how overconfidence influences the prediction quality and ultimate relevance of the outcome variables of information markets.

The results are synthesized in **Part 4**. Here, we specifically emphasize the integration of subject-level results from the first experiment and group-level results from the second. We discuss the empirical results in light of the experiments' particular limitations. After deducing the practical implications of these results for the innovation management discipline, we conclude this work by highlighting its implications for future research.

Part II.

Conceptual Background

2. The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty

This chapter introduces the reader to the difficulty and importance of valid innovation evaluation. At the same time, we aim to generate a more substantiated understanding of why a research focus on innovation evaluation via information markets is particularly promising.

In Section 2.1, we highlight important characteristics of innovations for our research. We then address uncertainty in the context of innovation in Section 2.2.1, beginning by exploring and discussing different concepts of uncertainty. From there, drivers of uncertainty in innovation development are identified and qualified in Section 2.2.2. Our introduction to uncertainty in the context of innovation concludes in Section 2.2.3, where we highlight the detrimental impact of uncertainty on activities that are related to innovation evaluation.

We ultimately address how uncertainty may be reduced so as to increase innovation evaluation quality in Section 2.3. Two particular dimensions for reducing uncertainty are explored, first in Section 2.3.1, where we identify relevant sources of information, and second in Section 2.3.2, where we evaluate mechanisms to aggregate and evaluate the information provided.

2.1. Relevant characteristics of innovations

The essence of innovation is the creation of something new. More than 60 years ago, Schumpeter refined the characteristics of innovation in his seminal work on business cycles, which is still reflected in much of today's definitional discourse (Schumpeter 1939). Schumpeter clearly distinguishes innovations from inventions:

Innovation is possible without anything we should identify as invention,

2. The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty

and invention does not necessarily induce innovation, but produces of itself [...] no economically relevant effect at all. (Schumpeter 1939, p. 84)

He regards inventions as uniquely novel problem-solving entities, while innovations are the product of a novel combination of production factors (Schumpeter 1939, p. 6). Innovations may incorporate inventions, but inventions are not critical precursors for innovations.

Schumpeter roots his definition in the abstract and in simplifying macro-economic variables (Ruttan 1959). However, innovation management scholars commonly distinguish between three perceptual dimensions to more appropriately define innovation: objective (*What is new?*), procedural (*Where does new start and end?*), and subjective (*For whom is it new?*). These dimensions are seen in the micro-economic context, such as in corporate projects or entrepreneurial undertakings (Hausschildt and Salomo 2007, p. 9). As discussed above, innovations aim to **benefit the innovator economically** (Roberts 2007). Today, the exploitation terminology has evolved from Schumpeterian profit-oriented exploitation. This transition started with the observation that **many innovations are created by users to satisfy unfulfilled needs** by innovating. Following the pioneering work of Von Hippel (2005, p. 177), innovation management scholars have widely acknowledged a broader understanding of exploiting an innovation's value, which also accounts for public and private need fulfillment without necessarily aiming to generate financial profits.

In sum, innovations are novel processes, physical objects, or any combination thereof and they need to be distinguished from inventions. Their features are not necessarily comprehensively new but may stem from a novel recombination of existing features and their *raison d'être* is profit- or need-oriented exploitation.

Independent of whom they are to benefit, innovations are regarded as investments in potential solutions to problems that are either currently unsolved or are expected to surface in the future. Profit-oriented companies engage in innovation to secure future cash flows. **Only innovation can ensure that companies continuously meet and serve customer needs** (Hauser et al. 2006). Empirical studies document that organizational growth and profit rates are positively affected by successful innovations (Bayus et al. 2003; Sorescu and Spanjol 2008).

However, existing research has shown that few innovation projects succeed. According to Stevens and Burley (2003), **between 40 and 75% of new products fail**, and a study by the consultancy Booz-Allen and Hamilton showed that no more than 25% of

innovation projects that enter development become commercially successful (Booz-Allen 1982). Such failure rates explain the profound need to validly evaluate the success potential of innovations and their underlying characteristics before allocating resources to their development (Stockstrom and Herstatt 2008). Proper evaluation allows companies to conceptualize, select, plan and execute innovation undertakings that align with their goals. However, the above-mentioned failure rates indicate the difficulty of evaluating the relevant characteristics and success potential of innovations (Kaplan et al. 2003; Denrell and Fang 2010).

Moreover, the difficulty of evaluating innovation success potential has frequently been related to an innovation’s *innovativeness* (Reid and De Brentani 2004). Here, a common theme **equates an innovation’s degree of innovativeness with its newness** (Garcia and Calantone 2002). The same subjective, objective, or procedural dimensions that define innovations, are now called upon, but rather than discriminating innovations from non-innovations in a binary fashion, the scale is formatted in a more fine-grained manner to identify how *innovative* or *new* an innovation really is. Scholars commonly choose subjective, market-oriented perspectives by comparing innovations to existing processes, products or services in the target market. This seems sensible based on the exploitation-oriented and environment-dependent nature of innovation (Hausschildt and Salomo 2007, p. 23).

Existing research has shown that the success of highly innovative, new products appears to be particularly difficult to evaluate (Reid and De Brentani 2004). Compared to more incremental innovations, highly innovative undertakings often yield higher failure rates during their development (Neff 2005). Yet at the same time, these innovations are attributed higher long-term investment returns once they are successfully introduced to the market (Cooper 1990; Sorescu and Spanjol 2008). Highly innovative developments that break with existing technological paradigms often outperform existing technologies when customer requirements shift to different sets of performance variables (Christensen and Bower 1996). Innovative products are more likely to enjoy *de facto* monopolies after their introduction because they can be better protected by patents, and are harder to imitate by competitors (Tirole 1988).

In short, evaluating innovation success potential is a crucial activity for profit-oriented companies. At the same time, achieving valid evaluations of innovations is difficult, and this is particularly true for highly innovative undertakings. Such difficulty, however, can be offset when a product achieves market-entry success, which may highlight the augmented value of valid evaluation in highly innovative endeavors. It is therefore im-

portant to foster a more substantiated understanding about the drivers that obstruct valid evaluation of innovation, a topic that will be addressed in the following section.

2.2. Uncertainty in innovation

Section 2.1 indicated that innovations are critical for the survival of companies and that evaluating innovation success potential is at the same time important and difficult. The goal of this section is to develop a thorough understanding of why it is difficult to evaluate the success potential of innovations. We introduce the concept of uncertainty to explain the difficulty in evaluating innovation success potential. The first subsection discusses different concepts, facets and degrees of uncertainty in the context of innovation. We then focus on factors that drive uncertainty in innovation endeavors. The final subsection discusses the impact of uncertainty on innovation-related decision making, with specific focus on the evaluation of innovation success potential in the early phases of innovation projects.

2.2.1. The concept of uncertainty in innovation

Research in information and decision theory has traditionally characterized uncertainty as a state in which different potential future outcomes have been identified but where the underlying probability distributions of these future outcomes remain unknown (Schrader et al. 1993; Brun et al. 2009).

Situations in which organizations evaluate the success potential of an innovation project often resemble such a state. Consider the roll of a dice as analogous to engaging in an innovation project: Only when we have a fair understanding beforehand that a dice roll scores a 6 at a probability of $\frac{1}{6}$ and that such a roll will pay out an expected amount x , can we properly evaluate the value of betting an amount z on that outcome. In this case, we would be facing a risk that could be perfectly quantified, since the underlying probability distributions are known. An innovation, however, resembles a roll of the dice where the characteristics of the dice and its environment are widely unknown *ex ante*. Imagine receiving the dice blindfolded by a stranger and having to roll it on an unknown surface. Under such conditions, nobody could properly assess the risk of betting on a 6 because he would lack information about the probability of that outcome. As Hurst (1982) puts it: “Innovation is a process where one steps into the unknown.”

Accordingly, and frequently cited by innovation management scholars (e.g. Tatikonda

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

and Rosenthal (2000); Herstatt et al. (2004); Lee and Veloso (2006)), Galbraith describes uncertainty in organizational tasks such as innovation undertakings as “the difference between the amount of information required to perform the task and the amount of information already possessed by the organization” (Galbraith 1973, p. 5). Here, information can be understood as verbally encoded knowledge (Glaser et al. 1983). His concept of uncertainty as focused on a lack of information is reflected by later approaches to defining uncertainty. For example, Brashers (2001) refers to states of uncertainty as being when details of situations are ambiguous, complex, or probabilistic; when information is unavailable or inconsistent; or when people feel insecure in their own state of knowledge or the state of knowledge in general.

Schrader et al. (1993) provide a more faceted perspective on uncertainty within problem solving processes such as innovation endeavors, extending the uncertainty concept in two ways.

First, they differentiate between **lack of information** (uncertainty) and **lack of clarity** (ambiguity) as distinct dimensions. Lack of information refers to the case in which the problem solver does not know the factual values of the variables deemed relevant to the problem. Consider, for example, an innovator who is convinced that the total number of potential customers for his innovation is an important variable but who does not know how many potential customers exist. According to the definition of Schrader et al. (1993), the innovator is exposed to uncertainty because he lacks information regarding the number of potential customers. The level of clarity describes the degree to which the problem solver is satisfied with his understanding of the problem structure and the underlying problem solving algorithm, e.g. the variables and variable relationships relevant to the problem. As an example of lack of clarity, the innovator may be ambiguous about whether he should consider different modes of distribution to estimate the number of potential customers.

Second, Schrader et al. (1993) argue that **lack of information and lack of clarity are not exogenous to the problem but endogenously chosen** during the problem framing process. The problem solver decides how uncertain and ambiguous he wants to render the problem at hand and thereby actively influences the potential solution space, the resources needed, and the appropriate organizational context. However, the authors also stress that the choice will often be made implicit, stemming from the problem solver’s preferences, experiences, educational background, and the capabilities, policies, and needs of his organization.

In conclusion, Schrader et al. (1993) infer that the problem solver (or innovator) plays

a crucial role in defining how much information and clarity are perceived to be missing at the beginning of an innovation task. They suggest that the innovation process will largely depend on this subjective definition.

This notion of uncertainty as an endogenous variable is supported by researchers who stress that uncertainty is a subjectively characterized state that depends on how a person assesses the probability of an event (Babrow 2001). A curve-linear function describes the relationship between uncertainty and beliefs about probability. Uncertainty is lowest when subjects believe that the probability of an event’s occurrence is 0% or 100%, and highest when the probability of occurrence is believed to be 50%.

Although uncertainty is commonly characterized as an undesirable state, subjects who initially embrace uncertainty may actually improve their decisions because they may reach an agreement when “honest differences in fact and values might otherwise lead to intransigence” (Hanft and Korper 1981). In fact, the willingness to accept uncertainty has been positively related to societies’ ability to generate innovations (Shane 1995). From an evolutionary perspective, uncertainty acts as a precursor for innovation because people can have different and conflicting beliefs that allows them to engage in competition and generate novel solutions (Foster 2010).

For the following work, we will define uncertainty as a state in which insufficient information is available about the set of variables, variable relationships, or variable characteristics that are relevant to effectively framing and solving an innovation task.

2.2.2. **Drivers of uncertainty in innovation**

The previous section introduced the concept of uncertainty and suggested that uncertainty is imperative to innovations, as they resemble problems with unknown outcome distributions. This section will discuss drivers of uncertainty in the context of innovation in order to generate a better understanding of *why* it is particularly difficult to access sufficient information in innovation endeavors.

Many researchers concur that uncertainty is greatest at the beginning of an innovation project (Koen et al. 2001; Reid and De Brentani 2004; Brun et al. 2009). At this stage, very little information exists about the set of relevant variables and their corresponding values in order to validly assess the risk of allocating resources to the innovation project (Montoya-Weiss and O’Driscoll 2000). As a result, the beginning of an innovation project is commonly referred to as the “fuzzy front-end” of innovation development (Koen et al. 2001). Herstatt et al. (2004) portray the development of an innovation in a five-phase

2. The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty

process, as depicted in Figure 2.1. In this process, the **fuzzy front-end** describes all activities in the first two phases. These activities cover generating and evaluating ideas, refining potentially successful ideas into more elaborate concepts, developing them, and introducing them to the market (Kim and Wilemon 2002).

Decisions at the fuzzy front-end are considered exceedingly important for the outcome

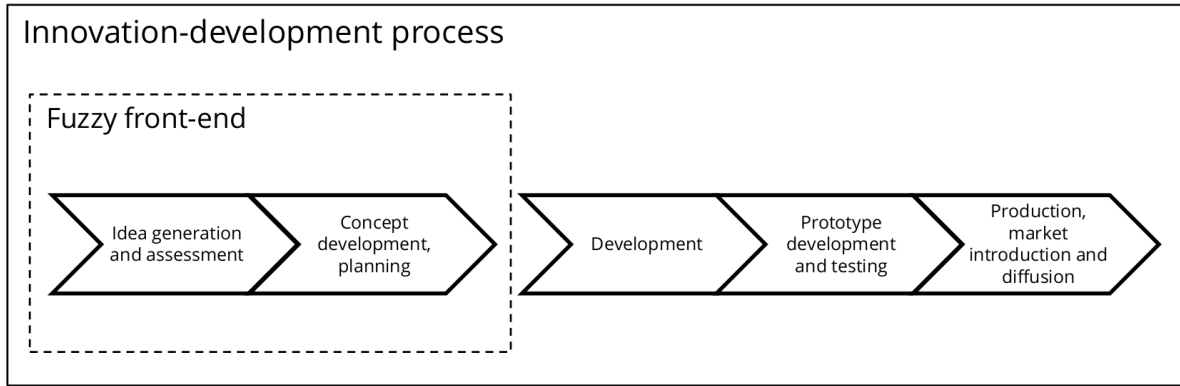


Figure 2.1.: The fuzzy front-end of the innovation process (Source: Own depiction based on Herstatt et al. (2004))

of an innovation project because they have a considerably large effect on all subsequent actions and investments (Cooper et al. 1998). Up to 75–85% of total product life cycle costs are determined during the idea selection phase, yet only 5–7% of the total costs have been generated at this point (Creese and Moore 1990). The underlying characteristics of novel products are strongly shaped in the early phases of an innovation project (Zhang and Doll 2001); failing to provide clear specifications at the outset often results in costly and timely delays during later development phases (Kim and Wilemon 2002). Cooper (1994) aptly summarizes that “the greatest differences between winners and losers [in innovation development can be] found in the quality of pre-development activities.”

Considering the high degree of uncertainty and the importance that is attributed to decisions at the fuzzy front-end, it appears sensible to **focus the drivers of uncertainty in the early stages of innovation development** (Brun et al. 2009). Here, the most may be gained by identifying and understanding the drivers of uncertainty, so as to prevent innovators from making decisions based on ill-informed evaluations (Smith and Reinertsen 1992; Reid and De Brentani 2004). Souder and Moenaert (1992) point

out that the front-end holds the greatest potential for improving the quality of decisions with the least possible effort (see also Stockstrom and Herstatt (2008)). The following paragraphs aim to identify those factors that particularly drive uncertainty at the fuzzy front-end of innovations.

Jalonen (2011) provides an extensive literature review on sources and drivers of uncertainty in the context of innovation. Based on his findings, we distinguish between external (or environmental) and internal drivers of uncertainty.

The **external drivers of uncertainty** can also be described as market-related drivers, i.e. those stemming from the innovation-related external market-environment. As stated earlier, the idea of innovation implies that it is being implemented to meet the needs of the market. Market-based drivers of uncertainty at the fuzzy front-end can be partitioned into the following categories: customer-driven, competitor-driven, supplier-driven and institutionally-driven uncertainties, as well as uncertainty that stems from the market's evolution or dynamics (Zhang and Doll 2001; Jalonen 2011).

The first and most important category covers uncertainty related to lack of information about potential **market needs and customer characteristics** (Souder and Moenaert 1992; Harris and Woolley 2009). It remains uncertain at the fuzzy-front end whether individual **customer needs** will translate into broader market needs because individuals' needs may be narrowly related to their personal interests (Enkel et al. 2005). Identifying needs based on a few customers drives uncertainty regarding the market potential of related solutions.

Even if no gap is believed to exist between stated customer needs and market needs, uncertainty can be fueled by a lack of information about the validity of the statements. Potential customers often fail to properly evaluate their own abilities to use innovative new products, even if they are allowed to briefly experience them. This stresses the uncertainty that stems from the lack of validity of customer feedback regarding innovative ideas. Customers exhibit a gap in perceived use value before and after experiencing an innovation. This is particularly true for highly innovative products because they lack similarity to existing solutions (Billeter et al. 2011).

Furthermore, defining price ranges drives uncertainty for innovative new products because uncertainty is not automatically curbed by competition from similar products (Christen 2005). Although properly measuring willingness to pay is a crucial aspect in evaluating new product ideas, managers often neglect to probe potential customers for

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

this information and purely cost-based pricing does nothing to reduce demand-side uncertainty (Bergstein and Estelami 2002).

Especially in business relationships, customer uncertainty regarding an innovation's use value may often obstruct its further development. Customers are more uncertain about the idea's value than the innovator because they possess even less information about the innovation's beneficial characteristics. Thus, potential customers may be easily offset by a well-defined requirement to change processes and sourcing if these incur higher set-up or production costs (Gassmann et al. 2010). The innovator remains uncertain whether negative customer feedback is based on similar factual characteristics or biased by customers' uncertainties regarding the innovation's value.

Second, lack of information about **the competition and its actions** drives innovation-related uncertainty. Some authors have argued that increasing globalization and market liberalization are important drivers of competition-related uncertainty (Jalonen 2011). The more easily potential competitors can access relevant target markets from anywhere in the world, the more uncertain innovators will be about the presence of potential competitors in these markets.

Uncertainty additionally results from missing information about the nature of competition in technologically novel or rapidly changing markets. This is particularly relevant in the context of innovation (Courtney et al. 1997). Consider, for example, the competitive changes that occurred in the music industry during the first decade of the current millennium. In 2000, mail-order and brick-and-mortar record stores were the sole distributors of music records. Today, music is additionally sold by completely new and highly different types of competitors, including streaming services that offer music as a service or online retailers that sell music solely as digital products like mp3 files or the right to listen to specific tracks online. In this competitive environment, the previously existing actors faced high degrees of uncertainty in how they should develop and differentiate their business models (Meisel and Sullivan 2002). This example furthermore illustrates the close relationship between uncertainty about competitor behavior and customer needs. Uncertainty stemming from competitor behavior appears to be positively related to uncertainty from changing customer needs or increasing access to innovative resources such as mp3 players or broadband internet connections.

Third, **regulatory and institutional conditions** can create uncertainty for innovating organizations. Companies that aim to enter existing markets with innovative products

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

often face uncertainty when meeting customer needs sufficiently indicates market success potential. New market entrants are not associated with existing institutional actors such as industry associations, or their market functions. At the same time, such actors are mainly staffed and organized by incumbent firms that are unable or unwilling to understand and support the marketing activities of new market entrants (Vermeulen et al. 2007).

In addition, complex regulatory environments often drive uncertainty because they increase the difficulty of identifying relevant product characteristics besides customer requirements, which has been shown for medical innovations in Japan (Numata et al. 2010) and consumer product packing innovations in Europe (Heiskanen et al. 2007).

Finally, and more closely related to the internal uncertainty regarding targeted solutions and their technological characteristics, innovators' uncertainty is positively influenced by a **lack of information about external resources** from suppliers and how suppliers may contribute innovation development. Especially when innovations require new types of materials, supply-related uncertainty is greatly increased. Such uncertainty may be further emphasized if new suppliers need to be found. New suppliers may only be able to provide insufficient information to assess the prices and quality of their products and services (Hoetker 2005).

The **internal drivers** of uncertainty comprise the lack of information about technology- and market-related organizational resources and capabilities, which may often be fueled by low quality in internal communications.

Lack of information about the fit between innovation opportunity requirements and organizational resources and capabilities is arguably the most important driver of internal uncertainty (Souder and Moenaert 1992). In our discussion, resources refer to stocks of available factors that are owned or controlled by the innovating organization, while capabilities are the organization's capacity to deploy resources (Amit and Shoemaker 1993). Especially at the beginning of an innovation project, the innovation's technical feasibility, functionality, and quality are at least partly unknown because current information only allows inferences about how resources and capabilities match current products and services (Leifer et al. 2001). As long as the technical details are undefined, organizations will be uncertain about their ability successfully develop the innovation. Furthermore, the less the innovation's technological characteristics resemble current products or services, the more uncertain the innovating organization will be.

Uncertainty additionally increases when communication quantity and quality is insufficient, as this prevents the flow of information that could deliver diagnostic information regarding the innovation to relevant recipients. Brun et al. (2009) refer to “multiplicity of the subject” when alternative meanings arise due to information from different reference points being provided. For example, uncertainty can be driven internally by organizational units that perceive situations differently when identifying and evaluating innovation opportunities, but that lack the means to communicate diagnostic information appropriately. In particular, the interface of market-oriented organizational units that command problem- or need-related information, and technical units that command solution-related knowledge, is a well-known driver of uncertainty (Hall et al. 2011). Marketing and R&D departments have garnered considerable notoriety for increasing uncertainty at the front-end of innovation by vehemently disagreeing over the new product’s preferred characteristics without exchanging sufficient information. Disputes are fueled by low relationship quality, difficult-to-explain domain knowledge and different environmental conditions, all of which can ultimately lead to decreased comprehension and credibility among organizational units (Moenaert and Souder 1996). **Social cohesion, communication and inter-functional coordination are frequently insufficient** to provide effective information exchange, which increases decision makers’ uncertainty about the success potential of underlying innovation projects (Souder and Chakrabarti 1978; Hise and O’Neal 1990).

In sum, we can conclude that uncertainty in innovation development is particularly large at the beginning of an innovation project. Uncertainty stems from external and internal dimensions. External dimensions describe the characteristics of the innovation’s target market, including customers, competitors, suppliers and market dynamics. Internal dimensions refer to the characteristics of the innovating organization’s resources and capabilities for successfully developing solutions that meet market needs.

2.2.3. **The impact of uncertainty on innovation-related decision making**

There exists a wide consensus in innovation management research that uncertainty has a strong influence on innovation-related decision making and outcomes (Tatikonda and Rosenthal 2000; Herstatt et al. 2004; Loch et al. 2008). Section 2.2.1 highlighted that uncertainty is a necessary precursor to innovation in that it creates (at least perceived)

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

opportunities to better meet market needs. The previous section discussed how internal and external drivers of innovation-related uncertainty often lead to a lack of information regarding the innovation object. We now aim to foster a better understanding of the impact of uncertainty on innovation-related decision making, with a particular focus on the fuzzy front-end of innovation. Only with this understanding can innovating organizations evaluate whether and how to address innovation-related uncertainty.

First, **uncertainty can cause significant delays** in decision making at the fuzzy front-end of innovation. Stockstrom and Herstatt (2008) cite a large scale German interview study by Bullinger (1990), which reports that one-third of all product-development efforts are unnecessary changes that prolong project completion times. Retrospectively, the interviewed corporate managers claimed that information to avoid these wasted efforts was often available at the time the innovation project was initiated. Uncertainty can increase delays in highly dynamic market environments in which market and technological conditions quickly change. Strong and frequent changes in the environment decrease the chances of effectively matching a given set of resources and capabilities to market needs and determining organizational requirements (Khurana and Rosenthal 1998). Organizations that are less flexible in allocating resources and unwilling to experiment and fail will be more likely to delay innovation development in situations of high environmental uncertainty (Eisenhardt and Tabrizi 1995).

Second and more importantly, **uncertainty has been closely related to the quality of evaluation during the course of innovation**, i.e. the higher the uncertainty, the lower the quality of innovation-related evaluation will be (Kim and Wilemon 2002; Brun et al. 2009).

High degrees of uncertainty have been positively related to avoiding the dealing with of that uncertainty. Organizations are more likely to be able to counter uncertainty if they are sufficiently informed about how much time and money they must invest to reduce particular dimensions of uncertainty. In cases where uncertainty is very high, organizations will often rely on existing information, which may foster ill-informed decisions. For example, uncertainty about the distribution of demand functions commonly leads companies to charge too little for a new product because they underestimate the innovation's additional value. Instead, they tend to focus on variable production costs to determine prices, and furthermore, neglect investments in the exclusive resources and capabilities that allowed the innovation's development (Marn et al. 2003). Similarly, uncertainty

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

about customers needs in changing market environments drive incumbent firms to rely on the feedback of existing customers. While this may decrease uncertainty about the demands of this particular group, it will often lead to faulty assessments if new customer groups (and needs) emerge and grow in importance (Christensen and Bower 1996).

When information gaps are still present after the evaluation process, they may prevent investment in new products even with low uncertainty regarding the potential customer benefits and internal ability to meet customer needs. A recent Japanese study of the medical industry found that uncertainty with regard to long-term commitments to regulatory processes during the innovation development process is a central obstacle that hinders innovations from moving from the concept- to the development-phase (Numata et al. 2010).

As a consequence, **uncertainty has been negatively associated with innovation success potential** (Stockstrom and Herstatt 2008).

For example, Souder (1988) found that uncertainty from lack of communication may negatively impact decision quality when selecting and pursuing innovation. In particular, disharmony between marketing and R&D departments appears to have a negative influence on innovation success potential. He concurs that disharmony prevents effective communication, i.e. the flow of information that can reduce uncertainty with regard to the need-solution fit. Decisions are not only delayed but are also based on insufficient levels of information, which ultimately leads to erroneous decisions. Tatikonda and Rosenthal (2000) studied the impact of task uncertainty in new product development projects within high-tech firms. They found that higher degrees of uncertainty due to technical novelty has a significantly negative influence on the project success variables of “time-to-market” and “unit-cost objective.” However, they also found that technical performance objectives are positively associated with uncertainty due to technical novelty in new products. Companies that engage in innovation with a high degree of technological uncertainty compromise new product performance by significantly overstressing development schedules and costs. The authors argue that high-tech firms underestimate development performance, and more importantly, overemphasize the achievement of technological goals compared to factors that are relevant to business success.

To conclude, uncertainty can have considerably negative impact on the speed, evaluation quality, and subsequent success of innovation projects. These findings strongly suggest that innovating organizations should actively engage in activities that help to reduce

uncertainty. Such activities may increase the quality of innovation evaluation, which appears to have positive effects on the success of an innovation undertaking.

2.3. Reducing uncertainty

It has so far been indicated that innovation success potential is negatively associated with uncertainty during innovation development. The previous sections identified drivers of uncertainty within innovation endeavors and illustrated the detrimental impact of uncertainty on the quality of innovation evaluation.

Now it is important identify and investigate means to help reduce uncertainty. Defining uncertainty as a state in which information and understanding are lacking supports the notion that uncertainty may be reduced by increasing the quality and quantity of relevant information (Jalonen 2011). Although gaining deterministic knowledge about absolute the likelihood of an innovation's success is not possible, sufficient effort should be devoted to reducing uncertainty during innovation development (Frishammar et al. 2011). While innovation managers must assess how much effort is *sufficient*, **research indicates that practitioners should engage more vigorously in reducing innovation-related uncertainty** (Enkel et al. 2005).

In his review of methods for evaluating innovation, Ozer (2005) indicates two particular dimensions that may effectively help innovators reduce uncertainty and increase innovation evaluation quality, as depicted in Figure 2.2. In the context of organizational research, Kumar et al. (1993) refer to the two dimensions as “*selection*” and “*agreement*” problems in gathering relevant organizational information. First, uncertainty can be reduced by **increasing the absolute amount of diagnostic information**. Innovators can only render decisions on the basis of valid information if that information has been successfully sourced prior to decision making. Second, the benefits of valid information can be extended by applying effective mechanisms by which to **filter and aggregate such information**. The more effectively information is filtered for random noise, the more valid the innovation evaluation will be. Figure 2.2 implies the multiplicative relationship between filtering effectiveness and amount of helpful information sources, which is supported by existing research. Zack (2001) argues that an organization's “ability to predict, infer or estimate” is positively influenced by the “organizational and technical resources and capabilities to locate [...] factual knowledge reliably and meaningfully.” The following sections will highlight both facets of uncertainty reduction.

2. The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty

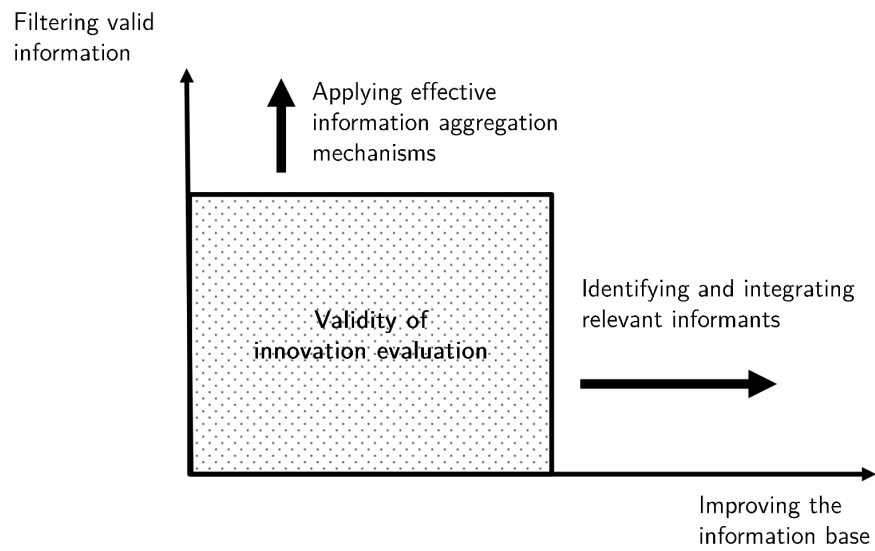


Figure 2.2.: Reducing uncertainty and increasing the validity of innovation evaluation by improving the information base and applying effective evaluation (Source: Own depiction)

2.3.1. Improving the information base

Uncertainty reduction starts with the collection of relevant information to improve the information base. Previous research has discussed how to effectively source information for innovation evaluation.

First, Armstrong (2001b) addresses the merits of **sourcing information via quantitative or qualitative approaches**.

Quantitative methods for sourcing information are commonly based on mathematical extrapolation. These methods provide functions that use quantitative past data to estimate the (future) potential of innovations. For example, the innovator may forecast the sales figure for an innovative new product by using the sales figures of a predecessor. Such extrapolation can provide promising results when evaluation objects and conditions are comparable to the predictors and sufficient data exists. This applies to stable relationships between product characteristics and market environment (Armstrong 2001a). However, Armstrong (2001b) also stresses that innovation environments are most likely unknown *ex ante* (i.e. the innovation has not previously interacted with the market), which precludes being able to make inferences about their stability. In uncertain environments where little unambiguous past data exists, forecasters have been advised to incorporate qualitative information to generate informed evaluations (Armstrong 2001b; Ozer 2005).

Second, the existing research discusses whether **the number of informants** is relevant to evaluation quality (Kumar et al. 1993). Van Bruggen et al. (2010) suggest that the key informant approach is most widely used in practice because it represents the most simplistic approach to sourcing information. Here, a single informant is selected because of his relative knowledge and willingness to share that knowledge. However, the authors emphasize that this approach is subject to significant drawbacks, including exposure to informants' biases, random error, and the inability to aggregate information from domains that are unavailable to the key informant. They concur with Armstrong (2001b) that sourcing information from single informants is inferior to using multiple informants if that data are effectively aggregated.

Empirical research supports the notion that the quality of innovation evaluation benefits from enlarging the breadth of information sources. However, only Leiponen and Helfat (2010) have directly studied the impact of breadth with regards to innovation objectives and information sources. Here, breadth refers to the amount different objec-

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

tives and sources. The authors found that broadening both dimensions helps companies achieve higher innovation success through more informed decision making. Based on their findings, they argue that (1) broad objectives help innovating companies to reduce uncertainty regarding the relevant relationships between innovation-related resources, capabilities and market variables and (2) broad information sourcing helps in identifying more effective objectives and better qualifying the relationships that are derived from the objectives.

The previous two paragraphs emphasized the value of sourcing qualitative information from multiple informants with heterogeneous backgrounds. It is now important to identify the **types of informants** who can contribute relevant qualitative information. Heterogeneous information, which organizations can use to improve forecasting or evaluation quality, can be located amongst different parties involved in the innovation process or amongst groups and individuals such as experts who possess transferable information from analogous situations.

Existing research from the domain of **environmental scanning** points to the dimensions of information sources that are relevant for the evaluation of innovation. Here, scanning describes the search for relevant information, where the context in which the scanning is performed is the environment. Originally stemming from strategic decision making on an organizational level, environmental scanning commonly describes the internal communication of information that is external to the organization (Albright 2004). In the context of innovation or new-product evaluation, however, environmental scanning can refer to any activity that aims to carry innovation-related information to the innovators, and which does not exclude sources of information internal to the organization. Ahituv et al. (1998) categorizes scanning sectors into innovation-task environment and general environment. **The innovation-task environment comprises information from the competitor, customer, and technological sectors.** The general environment consists of the regulatory, economic, and socio-cultural sectors. In their study, the authors focus on task-specific scanning in a sample of 40 Israeli firms that are split into successful and unsuccessful innovators. The results show that more successful innovators tap into written and personal external information from the task environment, whereas less successful innovators are more likely to rely on personal internal sources for information on competitors, customers and technology.

Klevorick et al. (1995) conducted a supra-industry survey to identify information sources

that are accessed by organizations in developing and evaluating new technologies. The study shows that manufacturing firms most commonly draw information from customers and suppliers; however, the authors add that external experts with very specific knowledge domains, such as university institutes, government agencies, market researchers, and other professional or technical societies, often provide information for identifying and evaluating technological opportunities that would otherwise be unavailable.

Tether (2002) relates the relevance of information sources to the internal and external dimensions of uncertainty introduced in Section 2.2.3. According to him, external informants comprise (potential) customers, competitors, suppliers and experts such as universities, consultants, and institutional organizations. He argues that reducing uncertainty from external and internal drivers necessarily requires the innovator to gather more information about the characteristics of these drivers. Companies that engage in more cooperative agreements when identifying and developing innovations are more likely to successfully develop highly innovative and complex innovations. This is because they source the amount of information required to solve uncertainty-related innovation problems.

Based on these findings, we focus our discussion on the more effective use of human subjects as informants to improve the information base in innovation evaluation. First, we will discuss the qualities of information that can be gained from the different informant groups introduced above. Then, we will discuss how the innovator can improve these contributions.

Potential and existing **customers** are an important information source because they ultimately determine the innovation's future use value, their willingness to pay for it and hence, future revenues. Existing research shows that seeking out customers to provide need- and solution-related information at the fuzzy front-end significantly improves the technical quality and speed of innovation projects (Carbonell et al. 2009).

The potential to gather innovation-related information from customers has been dramatically enhanced by extending the arsenal of techniques used to tap their knowledge. To reduce demand-related uncertainty, innovators need to learn about the formation of tastes and preferences, patterns of adoption, and the domestication of products considered social processes, which are deeply enmeshed in a variety of networks (Grabher et al. 2008). Over the last decade, the locus of these social processes has increasingly moved to an open domain that can be tapped by the innovator. Product-related virtual

2. The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty

communities (Füller et al. 2006), online product-recommendations (Granitz and Ward 1996) and open-source communities (Von Hippel 2007) allow innovating organizations to indirectly observe emerging customer needs and identify individuals who can contribute innovation-related information. This closes potential gaps between innovative product performance dimensions and customer needs, if the participants' needs truly resemble those of future customers.

Companies have also successfully recruited informants and allowed them to conceptualize and virtually construct new products ranging from micro-processors to LEGOs and Barbie dolls via user toolkits (Von Hippel and Katz 2002; Franke and Piller 2004; Piller and Walcher 2006). Employing user toolkits allows manufacturers to construct probability distributions of customer needs for specific product characteristics. As a consequence, innovators have started to delay moving the innovation from the fuzzy front-end phase to development until customer feedback has not only indicated market needs but until future customers have actually paid for the innovation up front. Crowd-funding platforms such as Kickstarter or Indiegogo have attracted considerable attention in recent times for reducing innovators' uncertainty regarding market needs, since they collect funds from customers who buy future innovations on the mere promise of the innovative idea or concept (Belleflamme et al. 2013). These customers ultimately provide highly diagnostic information regarding demand-side uncertainty.

However, customers may also provide biased information about the innovation's potential success. Christensen and Bower (1996) argue that established companies often fail to pursue emerging technologies because they do not currently meet existing customers' needs and can only be sold in markets that may appear unattractive compared to the current business. Yet, the authors also stress that such disruptive technologies may often improve and replace existing technologies at later time points, resulting in the loss of the very same customers who initially prevented the focal company from pursuing innovation in such a radical environment. These arguments highlight the importance of sourcing innovation-related information from a broad base that also encompasses potential customers, whose demands may precede technological shifts (Christensen and Bower 1996).

A small fraction of (potential) customers, also known as lead users, have been identified as particularly promising information sources because they have needs ahead of market trends and obtain significant benefits if these needs are fulfilled (Von Hippel 1986). Hence, the information they give has been regarded as highly diagnostic of the future demands of wider audiences. While lead users will often develop their own solutions to

meet their special needs (Lüthje 2004), they have been successfully used in the corporate context, where **they can provide valuable information to distinguish the potential of new product ideas, concepts and prototypes** (Lüthje 2000). They even act as information hubs that help to collect knowledge from potential future customers because this allows them to harness higher social status and community recognition (Von Hippel 2007; Jeppesen and Frederiksen 2006).

Employees hold valuable information about companies' innovation-related resources and capabilities and whether the conditions of the corporate environment foster innovation. While users may be more likely to hold market-related information regarding an innovation, employees play a crucial role in properly evaluating the innovation, thus increasing the viability of completing innovation endeavors. In other words, they may have less information about *what* an innovation should do, but are generally better informed about *how* it can be reached (Poetz and Schreier 2012). Employees exhibit exclusive technological, procedural and intellectual capacities that are particularly necessary in the evaluation of new product ideas and concepts (Amabile 1998).

We have already indicated in Section 2.2.2 that employees command very distinct types of knowledge and information. Marketing and R&D units are commonly attributed as having the most importance in the context of innovation evaluation because they are the most likely to possess need- and solution related information. While sourcing information from each unit individually may drive uncertainty, researchers have stressed the value of integrating both groups of employees when sourcing innovation-related information. As already discussed in Section 2.2.3, multiple informants help to increase the efficiency of transferring relevant knowledge between domains.

However, employees do not only contribute information regarding resources and capabilities. By sourcing information for innovation evaluation from staff, innovating organizations can also learn about organizational motivation and employee commitment to engage in innovation. Van de Vrande et al. (2009) conducted a survey among Dutch SMEs to learn about their motives for including heterogeneous informants in innovation evaluation. The study supports the notion that employees provide critical information about organizational ability and willingness to support innovation projects. The companies stated that resource-related and motivation-related are equally important when tapping employees as information sources in the evaluate of innovation success potential. Schweisfurth (2012) stresses the importance of having employees who are also innovative product users. Such “embedded lead users” can provide valuable insight regarding cus-

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

customer perspectives and facilitate information flows between users and companies during all phases of innovation development.

Knudsen (2007) focuses on **competitors** as valuable informants in the context of innovation evaluation and new product development success. She starts out by highlighting that inter-organizational collaboration in innovation should balance the advantages of accessing valuable, otherwise unavailable, information and sharing of risk and cost, against potential information outflows if relationships allow bi-directional knowledge exchange. Hence, it is important to restrict information flows, insofar as to prevent cannibalization of a company's own efforts. She shows empirically that drawing from competitors as knowledge sources during innovation development has no significant benefit overall, but may be helpful when competitors command supplementary information, i.e. information that is similar in codification but complementary in content. Such information likely minimizes the effect of cannibalization and can reduce market- and solution-related uncertainty, thus increasing innovation evaluation quality and innovation success potential. Some evidence exists that complementary information from collaborating competitors can create synergies in evaluating and developing innovations (Xu et al. 2013). These synergies are stronger when the underlying products are highly innovative because such an environment creates more potential for reducing uncertainty. Finally, Rindfleisch and Moorman (2001) add another important benefit of integrating competitors as information sources. In contrast to accessing information from customers or suppliers, overlapping knowledge and perceptions helps to reduce the danger of information getting lost in translation.

Moreover, **suppliers** have been identified as relevant informants in better evaluating the potential and characteristics of an innovation. Johnsen (2009) provide an extensive literature review that supports the notion that supplier involvement is particularly beneficial in new product development when they are closely integrated at early stages of development, such as during idea and concept evaluation. Clark (1989) studied the impact of supplier involvement in European, Japanese and US car companies during the fuzzy front-end and development stages of innovation projects. Controlling for product characteristics such as body type and price range, he found that supplier involvement can significantly reduce man-hours and development times. He argues that close involvement of suppliers in the early design phases allows them to effectively reveal relevant information. The car companies benefit from including suppliers' guest engineers because it

2. The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty

allows them to broadly source information with regard to industry developments and align the suppliers' R&D efforts with their companies' innovation goals. This reduces ill-informed innovation evaluation and planning, decreasing delays and increasing efficiency during the development phases. Including suppliers when evaluating innovations helps to prevent or reduce later design changes through early and intensive communication, and supports "first time right" development (Wynstra et al. 2001). Suppliers can help reduce uncertainty, shorten cycle times, and improve innovation quality upfront if they deliver information that helps determine material-solution relationships at the concept stage of innovation (Ragatz et al. 2002).

The final important group of informants comprises **technology and market experts**. Tether and Tajar (2008) surveyed the use of experts by 3996 UK-based companies when sourcing information for technological innovation activities. Of these, 26% replied that experts were very closely integrated into innovation development or were among the most important sources of information. The study shows that higher R&D activity and higher levels of innovativeness and technological complexity are positively associated with drawing on expert knowledge. The authors furthermore differentiate experts into business consultants, private research organizations (i.e. market research companies), and the public science base (i.e. universities). Management consultants are more highly involved in service-industry innovations such as finance and information technology, and both types of research institutes are more engaged in manufacturing innovations. Unfortunately, the article does not answer whether the expert information is provided in the context of innovation evaluation. Regardless, the researchers argue that expert knowledge can contribute ideas and insight by brokering knowledge from industries that are out of the innovators' scope. We may infer from this that such information would likely be helpful when evaluating the feasibility of technological solutions. Bessant and Rush (1995) aptly describes a technological consultant as analogous to a general medical practitioner whose main task is diagnosis (*Which innovation should the innovator pursue?*) and who then prescribes from a variety of available treatments (*What characteristics should the innovation have?*).

Thus, the innovating organization needs to **involve the informants** identified above in the organizational information search during innovation evaluation.

We have already emphasized that it is important to include informants with heterogeneous backgrounds, as their information contributes different perspectives to the evalua-

tion task. Furthermore, researchers in innovation management are increasingly arguing for **uniting informants from multiple domains** during the sourcing process to allow them to recombine knowledge to produce more innovative but also valid and relevant information (Alves et al. 2007).

Over the last decade, the paradigm of **open innovation** has gained increasing momentum. This idea assumes that firms can and should combine external and internal ideas and paths to market when developing innovations (Chesbrough 2003). Dahlander and Gann (2010) provide three important reasons why combining multiple internal and external actors in an open context is superior to more closed and internal approaches to innovation. First, social and economic changes in working patterns are reflected in open innovation approaches because they can better accommodate portfolio careers in which a job for life with a single employer loses importance. Individuals are better allowed to contribute to innovation independent of changes in their organizational home. Second, globalization has provided more opportunities to divide labor and access specialized skills more efficiently. Third, advances in communication technology have vastly increased the opportunity to store, access, and exchange existing knowledge across organizational and technological domains.

Current research presents findings on the impact of embracing open innovation in the context of developing and evaluating innovations. According to Laursen and Salter (2004), the larger the number of external sources of innovation, the more open firms will be in seeking external information and knowledge, since innovation often benefits from leveraging the insights of others. Dahlander and Gann (2010) highlight that resources will often become larger and more complex than a single organization can handle.

Though most of the research in the field has focused the benefits of opening up the innovation process (Chesbrough 2004), some researchers have started to identify and study the potential disadvantages. Limited cognitive ability of central actors in innovation evaluation processes may hinder the efficient search for and integration of external knowledge. Katila and Ahuja (2002) suggest that a curvi-linear relationship exists between innovation performance and the extent of external information acquisition, as too much information searching actually prevents firms from reaching conclusions and making decisions. Laursen and Salter (2006) present empirical support for this hypothesis in their survey of 2707 manufacturing firms in the UK. A wide and deep search for innovation-related information does not improve innovation performance indefinitely, eventually having a negative effect once a certain point is reached. Finally, many researchers have pointed out that certain industries find it much harder to internalize external informa-

tion with pecuniary effort, especially when technologies mature (Christensen et al. 2005) or industries are heavily influenced by intellectual property regulations such as in the pharmaceutical and information technology industries (Gassmann 2006b).

Overall, the increasing popularity of the open innovation paradigm appears to support the notion of integrating multiple informants from multiple domains when evaluating innovations; however, the disadvantages presented above support the importance of an effective brokering mechanism that allows the innovating entity to filter out invalid information, internalize external information and combine external and internal sources of valid information.

In sum, the information base upon which to evaluate innovations may be improved by two distinct means. First, emphasis should be placed on identifying and tapping sources of innovation-related information, and second, and valid information from varied groups of informants should be obtained.

2.3.2. Filtering valid information

The central goal of innovation evaluation is to decrease uncertainty when deciding about the initialization and characteristics of innovation endeavors so as to increase the likelihood of developing successful innovations. A large body of research has discussed the importance of applying meaningful methods for filtering data from multiple informants in the course of innovation evaluation (Ozer 2005). For information filtering to be effective, it is important to identify means to **aggregate responses into meaningful composite values** that can be used for decision making. Response data from informants with heterogeneous backgrounds very likely differ in type and quality (Van Bruggen et al. 2002). An expedient mechanism for aggregating information effectively incorporates the information that has been sourced from the informant groups presented in Section 2.3.1. We commonly find two distinct approaches to aggregating information from multiple informants, both of which will be introduced and discussed below (Van Bruggen et al. 2002; Van Bruggen et al. 2010).

Combined Judgmental Forecasts

Mechanistic aggregation, also known as combined judgmental forecasting (CJF), is used when informants do not interact or exchange information to form composite responses (Garthwaite et al. 2005). In CJF, all informants provide individual evaluations, which

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

are then aggregated into a group evaluation via a mathematical transformation. CJF covers a large array of methods that strongly vary in their mathematical complexity (Clemen 1989).

Arguably, the simplest approach to combining evaluations is to calculate unweighted group means from the individual information gathered. While unweighted group means may be comparably easy to calculate and provide protection against random errors in individual evaluations (Rousseau 1985), they do not prevent errors from systematic skew in judgment. If all individual reports underestimate the true underlying figure, averaging their responses will still yield an underestimation (Snizek and Henry 1989). As such, several researchers have proposed more elaborate methods to mathematically aggregate individual responses (Clemen 1989; Ozer 2005; Van Bruggen et al. 2002). For example, Van Bruggen et al. (2002) introduced (1) response-data-based weighted means and (2) confidence-based weighted means. The first method weighs the responses of agreeing informants more strongly than those of disagreeing informants. The second method weighs the informants' responses based on the informants' confidence in their responses. The authors compare these two novel approaches with in evaluating the future brand value of novel companies in a business simulation game. The results show that confidence-based means can significantly outperform unweighted, or response-data weighted evaluations. However, the authors stress that the task was perceived similarly difficult by all participants, which may have prevented systematic biases in the confidence-based weighting they applied.

Von Winterfeldt and Edwards (1986) emphasize that more complex methods, such as estimating Bayesian models (Agnew 1985) or complicated likelihood functions (Clemen and Winkler 1993) to mathematically aggregate individual responses, only improve the odds "[...] that you will simply be wasting your efforts."

Ozer (2005) discusses application areas for CJF with different complexity levels by drawing from a set of examples where CJF was used. Mixed results were found.

CJF yields satisfactory results when evaluation dimensions are well defined and similarly perceived by the informants. For example, Ozer (2005) refers to a case study in which a medical company aimed to evaluate the perceived importance of new product characteristics in medical devices. The device manufacturer aggregated ratings by doctors and nurses to effectively allocate R&D efforts (Ulwick 2002). Similarly, a case study of a Japanese software developer reports that mathematically aggregating the software-development teams' individual responses with regard to functionality and usability of innovative development environments proved useful in selecting an appropriate environ-

ment. The researchers argue that mathematical aggregation was feasible because the team agreed on a limited set of ten target criteria that were well understood and similarly perceived by all members (Miyoshi and Azuma 1993).

Still, the integration of CJF may fail when evaluating the potential and feasibility of new products. Two examples highlight reasons why CJF is particularly difficult to implement when evaluating innovations. Loch et al. (2001) developed a model for project selection at the German car company BMW. Their model aimed to evaluate the potential of 80 innovation projects based on 41 underlying criteria. Engineers first evaluated the criteria values and the importance of each, and these were then used to estimate a linear program to identify the most promising projects. This model was ultimately not used by the organization because the setup consumed too many resources and the model did not accommodate decision makers' need for quickness, ease, robustness and (graphical) transparency of results. In the end, it was too complex. Felli et al. (2000) underline how CJF for evaluating the potential of innovative projects may fail if informants become overwhelmed with the dimensions of information they are required to provide. In their case study, the authors develop a mathematical model that integrates multiple aggregation functions with the goal of selecting innovation projects for the Monterrey Bay Aquarium. Similar to the previous case, the model was ultimately not implemented because participants felt unable to provide all the values required. While successful evaluation was not completed via the model, the researchers agree with Loch et al. (2001) that the process of building the model benefited the collaborating organization in that they could better understand the important variables and variable relationships for selecting potentially successful innovations.

Although his examples do not provide comparisons with non-mathematical methods for aggregating response data from multiple informants, Ozer (2005) concludes that CJF presumably works best when clear and measurable performance indicators can be defined and understood before the evaluation. He points out that this may not be feasible in most innovation contexts because it is simply impossible or would require substantial resources.

Interactive Group Methods

Compared to CJF, an interactive group method (IGM) for innovation evaluation adds an important feature: Informants' data is now aggregated in an interactive process, meaning that informants learn about other informants' responses and may update their

own contributions after learning about these responses.

The characteristics of IGMs differ and depend on whether they are carried out via face-to-face interaction or remotely via electronic networks, whether or not informants are anonymous, whether participants share information verbally or through group support systems, and how information aggregation is organized (DeSanctis and Gallupe 1987).

By confronting informants with other participants' information, IGMs aim to create a consensual evaluation that is consciously shared by the informants (Rowe 1992). This is also considered the greatest challenge of applying IGM; it is often particularly difficult to reach consensus among participants with heterogeneous information or backgrounds. In CJF, participants cannot form a consensus because the aggregate is automatically generated before informants learn about results. When a consensus is reached in IGM, however, it is often be driven by power and personality rather than the quality of information, especially when information exchange is not carried out anonymously (Van Bruggen et al. 2010). Even in the case that power and personality can be neglected, groups may focus on information that was held in common before evaluation and which supports existing expectations and beliefs (Stasser and Titus 1985), thus reducing the group's ability to truly benefit from its heterogeneity.

Even still, several researchers highlight the advantages of requiring participants to interactively engage with information from distinct sources such as different organizational units when aggregating information.

IGM allows inter-domain communication, which can foster learning and span informational boundaries between the evaluating parties (Bonabeau 2009). Learning about different perspectives helps evaluators unmask and overcome private misconceptions. In contrast to CJF, participants receive the opportunity to reflect on their own responses in the light of other participants' information, which may reveal to them relevant relationships that would have otherwise been missed. For example, marketers may favor a certain technological idea because it has received superior customer feedback from a demo video, but might only learn through hearing engineers' responses that pursuing this particular idea would imply a steep increase in maintenance costs for potential customers. This novel information might therefore lead them to reconsider their initial evaluation and allow them to gather novel feedback from customers to update and improve their existing knowledge.

Zigurs and Buckland (1998) theorize that IGM evaluations of innovation-related problems such as judgment or fuzzy tasks lead to the best performance when information processes and communication support are particularly emphasized. This is because

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

transparent rules and rich communication are the most effective measures against conflict. In the same vein, (Souder and Moenaert 1992) argue that access to extra-functional information allows evaluators to reduce the variability of tasks and increase their analyzability, which may then free resources from innovation evaluation for use in innovation development.

Several case studies have emphasized the potential benefits of IGM in new product evaluations.

Several companies have employed IGM because they identified a major benefit in creating a common understanding about assessment criteria before evaluating the potential of new product ideas and concepts. For example, the electronics companies Hewlett-Packard and Ericsson used IGM to (1) develop a consensual set of key factors to look for in a new product and (2) allow the evaluation method to accommodate constantly evolving market environments, which are often present in the context of highly innovative products (Englund and Graham 1999). Moreover, the UK pharmaceuticals company ICI employed IGM in having project managers assess the potential of several R&D projects. In the related case study, Islei et al. (1991) found that mutual identification and evaluation of relevant criteria helped senior management to create a sense of ownership and identification with the underlying R&D projects, which ultimately had a positive impact on project outcomes.

In conclusion, IGM may be more suited for innovation evaluation tasks that require evaluators to identify and clarify relevant criteria for evaluation and benefit from inter-functional communication during the evaluation process.

The preceding paragraphs introduced **CJF** and **IGM** as popular methods for evaluating innovations through the aggregation of information from multiple informants. Based on previous research, we discussed the important benefits and pitfalls of both methods and illustrated their application in the context of existing case studies. With our work, we will particularly focus on IGM and will substantiate this choice by comparing the previous performance of both methods, as it relates to the context of innovation evaluation. In a recent study, Van Bruggen et al. (2010) experimentally compared the forecasting quality of CJF and IGM in two experiments. The researchers chose two kinds of forecasting tasks that were differentiated in their degree of information heterogeneity, or the variation in information into which a group of evaluators can tap. In low information heterogeneity situations, informants have access to similar information. This would be the case with a group of sales representatives for a regional sales forecasting task. In

high information heterogeneity situations, informants can tap into very different sources of information. One example would be marketing and R&D personal in a new product evaluation. The experiments show that IGM does not outperform CJF in situations where informants tap into common pools of information. When information is heterogeneously dispersed, however, IGM provides significantly better evaluations than CFJ. The authors attribute the relative performance gain in IGM to two intertemporal advantages over CJF: first, participants can improve their knowledge through the information exchange mechanism and mitigate the contributions of initially ill-informed participants; and second, contribution weights in CJF are typically taken at the beginning of the evaluation task. If they are not updated over time, weights may become inappropriate for the evolving and unstable environments that characterize high information heterogeneity situations.

By reviewing applications of IGM and CJF, Ozer (2005) arrives at a similar conclusion. He summarizes that **IGM may be more appropriate in uncertain innovation-related environments than mathematical models (such as CJF)**. He stresses that the quality of innovation evaluation benefits particularly from the interaction between informants with heterogeneous backgrounds and information (i.e. all parties involved in the innovation endeavor).

Characteristics and prerequisites of effective IGF

In the following paragraphs, focus is placed upon the prerequisites and characteristics of IGM that positively impact validity when evaluating innovative new-product ideas or concepts. We draw from the current body of research to identify and discuss characteristics of IGM that drive the quality of innovation evaluation.

First, the evaluation mechanisms' effectiveness appears to be positively related to its **openness and accessibility** to participants holding valuable and heterogeneous information, as it allows them to reveal information, learn from other information, and update their own views accordingly (Spann and Skiera 2003a; Van Bruggen et al. 2010; Rowe and Wright 2011). Increasing openness and accessibility allows increased exposure to valid information. This supports, in a straightforward manner, the information base, but also allows participants to better reflect upon their existing information. With each new participant contributing pieces of heterogeneous information, all receive more op-

portunities to acquire, compare, and improve individually-held information.

Existing research supports the notion of a positive relationship between openness and quality of innovation evaluation, and furthermore qualifies the dimensions of openness. Cowgill et al. (2008) studied the application of IGM for evaluating the potential of new services at a large IT company. The study particularly focused on the impact of **physical proximity** on evaluation outcomes. In their field study, the authors applied IGM on a corporate level, with participants from offices all over the world and spanning all functional units. By studying the correlation of the revealed information, the researchers found that physical proximity (sitting on the same floor) and cultural proximity (speaking the same Non-English native tongue) increased the positive correlation amongst participants' responses. Considering the benefits of incorporating heterogeneous information, it thus appears desirable to allow participation from multiple sites, independent of physical proximity.

Moreover, continuous and instant accessibility have been considered important drivers of evaluation quality via IGM (Van Bruggen et al. 2002), with **time-related openness** emphasized as an important dimension. We previously saw that uncertainty in innovation-related contexts often stems from rapid environmental changes such as market needs or alternative solutions. If informants are confronted with small time windows to reveal information or long processing times before information is released, they might be prohibited from revealing and learning about relevant information. While this could negatively impact evaluation quality *per se*, it might increasingly do so in very dynamic innovation environments.

Physical and time-related access restriction can be removed by establishing means to virtually access IGMs (Spann 2002). The last decade has given rise to extensive opportunities to connect individuals across geographical and organizational boundaries via virtual networks. IGMs can benefit from increasing inter-connectivity in order to provide instant and continuous access to information by any potential participant.

Informants must furthermore be **cognitively and organizationally empowered to access IGMs**.

Cognitive empowerment refers to ensuring informants' understanding of the IGM process so that their intent to reveal valid information can be carried through in their actions. Soukhoroukova et al. (2012) studied the application of IGM via virtual networks at a large German industrial company. The application was perceived well by the organization for assessing innovation potential. However, the researchers found that a main barrier to participation was lack of understanding about how to participate. In addition,

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

participants from different functional domains need to be able to absorb as much of the information that has been revealed via the IGM as possible, emphasizing the importance of translating information so that it can be commonly understood.

Organizational empowerment describes the mechanism's ability to accommodate a heterogeneous group of informants from distinct functional domains, hierarchies and organizations, and allow them to simultaneously engage in revealing and updating information. For example, we cited earlier that different functional units often fail to communicate effectively because they have tacit knowledge, or information that is specifically codified to their domain. Thus, the IGM needs to translate such domain-specific information in order to allow cross-boundary learning (Van Bruggen et al. 2010). In addition, uniting different hierarchies may negatively impact willingness to reveal information. Subordinates may refrain from revealing information in the fear that their information would shed negative light on their managers, and that thus, they might experience retaliation. As a consequence, IGM contribution intensity might be positively related to organizational rank (Garthwaite et al. 2005).

Second, the evaluation mechanism's effectiveness is increased when **appropriate incentives** induce participants to align information seeking and revelation with the innovator's goal of reducing uncertainty and false decisions (Wolfers and Zitzewitz 2004). Ostrover (2005) lists two goal dimensions that incentives in IGM should aim to achieve in the context of evaluation or forecasting tasks.

First, participants should be encouraged to seek, acquire and update relevant information before and during participation. Information search and updating must be incentivized because it creates search costs. While search costs can be low if participants draw from memory to gather information, they will be considerably higher if participants are required to engage in physical processes to retrieve information, such as studying documents, contacting third parties or conducting any form of research. Furthermore, information updating will often incur cognitive load for participants. Changing existing beliefs by updating previously held information is costly because it requires subjects to write down previously gathered information. Much research has documented that subjects are much less likely to absorb new information that questions existing beliefs than new information that supports those beliefs (Samuelson and Zeckhauser 1988; Ritov and Brainerd 1992). Thus, incentives are used to positively influence participants' willingness to assess and honestly integrate new pieces of information, even if that information contradicts previously held beliefs.

The second goal for IGM incentives is that they need to motivate participants to reveal information in an honest, timely and appropriate manner, in relation to the validity of other participants' information.

Participants face the cost of the time invested in signing up for and participating in the IGM. Therefore the IGM needs to provide incentives so that participants actually invest that time and contribute to higher evaluation quality by revealing helpful information. The method needs to attract participants who contribute valuable information and at the same time discourage participation from subjects who reveal flawed information. Incentives additionally need to prevent cases in which informants are motivated to withhold or reveal manipulated information, e.g. where information would shed negative light on superiors or reveal knowledge that was exclusively held by competitors.

Furthermore, the IGM should encourage participants to engage continuously and to reveal changes in private information quickly, as the method needs to rapidly absorb information changes to keep up with continuously evolving external innovation environments (Rothaermel and Hess 2007; Graefe 2009). Only then can other participants effectively update their information.

Finally, with regard to appropriate incentives, the quality of aggregated information may increase if the intensity of individual information revelation is sensitive to belief strength. Research shows that belief strength or confidence in personal information can be positively related to its validity (Van Bruggen et al. 2002). Consequently, IGM output may benefit if participants have appropriate incentives to communicate their information based on how strongly they believe in it (Kumar et al. 1993). This may temper truthful contribution to the existing body of information and other participants' inclination to draw from such information.

Ultimately, the IGM requires an effective **algorithm to coordinate the interaction of participants** in information exchange and revelation, **and to compute information aggregates** into meaningful and valid composite values (Spann and Skiera 2003a). Private information needs to be cheaply and validly accessible to the aggregation mechanism, and aggregated information needs to be effectively distributed amongst participants so that they can process and learn from that data. Final results need to be similarly understandable to the outside initiators of the IGM. After all, the goal is to yield information from human informants with heterogeneous backgrounds that has been filtered for validity and compressed via aggregation.

2. The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty

		Face-to-Face Meetings	Delphi and NGT	Information Markets
Openness and Accessibility	Participation without physical proximity	Very limited	Yes (for Delphi)	Yes
	Maximum number of informants	Very low	Low	Very high
	Cognitive barriers to participation	Low	Medium	High
Goal-compatible Incentives	Incentives for quick information revelation	Not formalized	No	Yes
	Incentives for valid information revelation	Not formalized	Possible but barely discussed in current research	Yes
Effective Algorithm for Aggregating Information	Continuous participation	Short-term only	No	Yes
	Sensitivity to belief strength	Not formalized	Yes	Yes
	Inter-informant learning	Yes	Yes	Yes

Table 2.1.: Accessibility and goal-compatibility fit for FTF, Delphi and NGT, and information markets (Source: Own depiction)

Graefe (2009) discusses the merits of three popular IGM types for aggregating information in the context of new product development: **Face-to-face Meetings (FTFs)**, **Delphi and Nominal Group Techniques (NGTs)**, and **information markets**. We will introduce each of them briefly and compare their fit with the prerequisites and characteristics for valid evaluation introduced above.

FTFs are arguably the most established and common form of group-based interaction for aggregating and exchanging information (Hiltz et al. 1986). They are unstructured in nature and require the synchronous presence of all participants. FTFs are traditionally carried out with all participants physically present in the same location but several technological solutions have been developed that allow virtual FTFs via computer networks. In an FTF, information is mainly shared by voice but vocal contributions will often be complemented by different forms of media such as computer presentations or hand-outs. Several factors impact the openness and accessibility of FTFs to heterogeneous informants in the context of innovation evaluation. FTFs exhibit natural limits in their physical and time-related openness to participants. First, FTFs are usually carried out in single sessions and take less than a few hours, which largely prevents the continuous integration of new information to the group's information base. This is because information can only be revealed and aggregated if it was present before or emerges during the

2. The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty

group session, which is less likely if session duration is short. Second, participants must be physically or virtually present during the FTF. On the one hand, physical presence may incur high logistics and opportunity costs if informants need to travel considerable distances. On the other hand, virtual presence may be hard to achieve in case the participants are situated in distant time zones.

Moreover, organizational factors likely limit the accessibility and openness of FTFs. As soon as individuals from multiple functional domains participate, requirements could quickly rise if domain-specific understanding hinders communication. Empirical evidence supports the notion that informants' unfamiliarity with each other likely has a negative impact on evaluation quality via FTFs when information is heterogeneously distributed. Gruenfeld et al. (1996) conducted a laboratory experiment that studied criminal suspect choices in FTF group discussions. The experimental design introduced different degrees of group member familiarity and information heterogeneity. If group members were familiar with each other before the discussion, they were particularly good at identifying the right suspect in cases where information was heterogeneously distributed. Unfamiliar groups only excelled in cases where members held comparable and full sets of information. Researchers relate situations in which aggregated group information cannot reflect all available information to social factors in FTFs that restrict openness and accessibility. Groups might not consider all problem dimensions if particular individuals dominate group thinking (Dalkey and Helmer 1963). Hierarchical and departmental dependencies often negatively influence participants' motivation to reveal information, since group members are not engaging in information exchange anonymously (Schütz and Bloch 2006).

However, FTFs may provide comparatively high satisfaction levels because individuals enjoy human interaction, which may foster inter-participant learning (Van De and Delbecq 1971).

Finally, the unstructured nature of FTFs may negatively influence the initiator's ability to coordinate and compute the aggregation of information. In an FTF, information can only be revealed by one person at a time, which limits the speed by which overall information can be increased.

Delphi and NGT introduce more structured processes to innovation evaluation. In NGT, individuals first generate independent evaluations. They then reveal their independent evaluations and enter an unstructured discussion to challenge individual reasoning. Finally, they retreat to form final individual evaluations and all individual evaluations

2. *The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty*

are averaged over the total group. NGT aims to intensify individual reflection on private and public information.

Delphi builds on the NGT approach but completely cancels out face-to-face interaction. The original Delphi method was developed by the RAND Corporation in the 1950's and has since been mainly used for policy evaluation and decision impact in public and corporate environments. Participants submit evaluations along with written explanations of how they reached their conclusions and how confident they are in them. Evaluations are then aggregated and participants receive feedback on group evaluations and explanations. This process is reiterated over multiple rounds until the process is concluded. The goal of Delphi is to provide iterated anonymous individual responses by well-selected experts that can be statistically analyzed on a group level (Goodman 1987).

Delphi traditionally requires a limited selection of experts as participants. Panelists in a Delphi process are supposed to learn from the others' expectations, which requires a sufficient ability on the part of participants to express their reasoning so that the other participants can understand it. Additionally, the number of participants is usually limited in order to reduce the information overload via abundant written explanations, which further limits Delphi's openness and accessibility (Green et al. 2007). Therefore, researchers stress the importance of carefully selecting the participants in a Delphi process to reflect all relevant information domains (Woudenberg 1991). However, Delphi allows participants to interact anonymously and virtually over a longer period of time, which increases accessibility across physical boundaries and allows subjects to give honest but unwanted feedback without fear of retribution.

With regard to appropriate incentives, Delphi participants usually receive incentives based on the overall prediction error of the group prediction. Delphi does not provide direct individual incentives for revealing new information. It may therefore remain questionable as to why or how individuals are motivated to invest effort in providing valid predictions if their individual contributions remain anonymous and unrewarded on an individual basis. The initiator may increase motivation to provide evaluation by recruiting experts who individually benefit from valid Delphi outcomes, thus reducing the necessity of providing rewards.

The Delphi method does well at aggregating individual opinions and reaching consensual judgments; however, previous research has argued that it may exaggerate consensus-seeking (Goodman 1987). Final group judgments may lack relevant individual evaluations that were unnecessarily aborted because of a tendency to conform to the majority view. As a consequence, Scheibe et al. (1975) argue that the stability of specific (sub-

)group responses over time should receive particular attention, in contrast to apparent consensus at the end of Delphi evaluations.

On the one hand, the availability of varying evaluations and supporting arguments from previous rounds allows the initiators to look beyond the ultimate group consensus and take into account deflecting arguments. The well-defined structure of information aggregation eases ex-post analysis of the aggregation process. On the other hand, information aggregation by rounds in fixed time intervals prevents participants from revealing important information in an ad hoc manner so as to increase the aggregation efficiency.

Information markets take a different approach by applying the principle of stock markets to forecasting. Information markets allow virtual trading on the outcome of future events by creating derivatives whose payout is tied to the future events' outcomes. Participants trade the derivatives or information market stock according to the outcomes they expect. The information market stock prices ultimately reflect the participants' aggregate expected value of the future outcome of an event, e.g. the launch of an innovation.

By relying on a market mechanism to aggregate information, information markets can effectively host a very large number of participants compared to Delphi or FTFs. Also, the market mechanism does not require any type of joint physical presence but can be accessed anytime and from anywhere, as long as participants are digitally linked. As information markets are usually continuously open for trading during their running time, participants can reveal updated expectations as soon as they receive and process novel information. While this makes information markets very open and accessible for participants, they are arguably the most difficult mechanism to engage in from the participants' perspective because understanding the principles of financial asset markets is required. Compared to FTFs, NGT and Delphi, information markets' physical accessibility can be offset by the cognitive barriers to participation (Green et al. 2007). Yet once subjects have understood the principle of information markets, the pricing mechanism prevents them from misunderstanding other participants' information, which could likely occur in FTFs, NGT and Delphi. As market prices develop, participants can update their beliefs based on market price changes and hence learn from other participants. Trading on the market usually remains anonymous and does not directly reveal motivations about why the market price (and thus the group prediction) changes. While changing stock prices do not transmit arguments and reasoning via trading, they provide unambiguously direct signals to fellow participants.

2. The Importance of Innovation Evaluation for Reducing Innovation-Related Uncertainty

Information markets are considered have incentives for the revelation of valid information (Arrow et al. 2008). Participants are incentivized to join and participate if they believe they possess information that will allow them to benefit from trading. This is deemed the largest benefit of using information markets to evaluating the outcome of future events, as compared to the previously introduced mechanisms (Spann and Skiera 2003a).

Information markets aggregate and codify group expectations via market prices. Additionally, participants' belief strength is reflected by the amount of stocks they sell or purchase. Based on research regarding the price efficiency of financial markets (Fama 1970), the aggregation mechanism of information markets has been deemed highly efficient in comparison to other IGFs (Arrow et al. 2008), as it provides compatible incentives for participation and information revelation.

Graefe (2009) compared all three types of mechanisms empirically and found that the Delphi method and information markets perform equally well, and both outperform FTFs. While subjects report that information markets are the hardest to understand, the author hypothesizes that they are “ [...] particularly valuable in situations where new information becomes continuously available and a large and heterogeneous number of participants have valid insight into the issue in question.” Yet interestingly, his experiments did not consider these values when comparing the mechanisms' effectiveness, which may have undermined the relative advantage of information markets in his studies.

A literature review by Bothos et al. (2009) on the research regarding information markets finds support for the open hypothesis by Graefe (2009). Information markets have been particularly chosen for evaluation tasks where information from large numbers of informants with heterogeneous backgrounds was required. This fits the information environment for innovation evaluation as described in Section 2.3.1, which may underline the appropriateness of information markets in an innovation-evaluation context.

In short, **IGMs appear most effective in the context of innovation evaluation.** Current research documents that information markets may be particularly promising in environments where a large number of participants have heterogeneous expectations, are geographically dispersed, and continuously yet asynchronously learn from new information (Graefe 2009). The previous decade produced a large strand of research that stresses the benefits of market-based aggregation mechanisms in providing incentives

for truthful and instantaneous information sharing, which are usually not provided in FTFs, NGT, and Delphi (Bothos et al. 2009; Dahan et al. 2010; Blohm et al. 2011). We build on these findings to **focus on information markets during the course of this study**. Chapter 3 offers a more thorough introduction to the application of information markets to innovation evaluation.

3. Information Markets for Innovation Evaluation

In the introduction, we indicated the promise of using information markets in the context of innovation evaluation. Information markets provide a novel approach for simultaneously sourcing and filtering innovation-related information. In this chapter, a more thorough introduction will be provided regarding information markets in the context of innovation evaluation. First, the fundamental mechanisms, theoretical foundations and primary applications are discussed in Section 3.1. Prior information market applications in innovation evaluation, their prediction objects and designs and outcomes are then illustrated in Section 3.2. Afterwards, we discuss different design choices in Section 3.3 and highlight their relationship with evaluation outcomes in the context of innovation.

3.1. Foundations of information markets

Information markets provide predictions for outcomes of future events. After defining prediction targets (e.g. the market share of an innovative product in a specific target market), the market's initiator creates derivatives called **information-market stocks**. The inherent values of these stocks are tied to the predictions' outcomes. A stock's value derives from the market share of the innovative product at time point t . Stocks are paid out according to the outcome immediately after t . For example, if the market share of the innovation is 40% at the moment the market closes, the stocks may be cashed out at 40 units of market currency.

Before the information market starts, participants acquire or are endowed with a portfolio of information market stocks and information market currency to trade the stocks. They are informed that their remuneration for participating will depend on their final portfolio value after the information market stocks have been paid out; the higher their final portfolio value, the higher their expected remuneration will be. Accordingly, if market participants hold different beliefs about the outcome of an innovation evaluation

3. Information Markets for Innovation Evaluation

issue, such as its likelihood of market success, they have an incentive to engage in trading. The person who believes that the innovation holds high market success potential allocates higher value to acquiring stocks that derive from the innovation's market success than the other participants. Accordingly, this person will try to buy stocks from the participants in exchange for information currency as long the person believes that the stock is undervalued and has sufficient amounts of cash. Over time, individual expectations regarding the prediction task's outcome are aggregated via the trading mechanism. Whenever participants acquire new information, they may form new expectations, which should incline them to buy or sell stocks and thus alter the information market's current prediction. Thus, incentives for participants are compatible with the initiator's goal of having market prices reflect participants' expectations regarding the information market stock's true underlying value.

Figure 3.1 presents a simple example to illustrate the basic principles of information markets. One should be aware that a real information market would very likely feature a larger number of traders and prediction targets.

1. At the beginning, the future events in question require precise definition so that their underlying prediction object is fully understood by potential participants. Participants' and initiators' expectations must reflect the same underlying prediction object prior to, during, and after the market, in order for incentives to align.
2. Next, participants with heterogeneous expectations and information decide to join the market. Because they believe that they possess superior and exclusive knowledge regarding the underlying prediction task, they expect to benefit from revealing personal information via their trading.
3. Before the information market starts, participants are endowed with an initial portfolio that consists of stocks and information market currency. In our example, there is one stock type and the currency is called virtual Dollar (v\$).
4. In the illustrated example, the expectations of the two participants differ. This opens the possibility for trading. The participants engage in trading via a continuous double auction mechanism (CDA): Participant 2 places a "sell"-offer at 35 virtual Dollars, which is absorbed by Participant 1's offer to buy stocks at prices up to 40 virtual Dollars. Accordingly, Participant 2 sells all her stock to Participant 1 for 35 virtual Dollars, from which the new portfolios and a market prediction

3. Information Markets for Innovation Evaluation

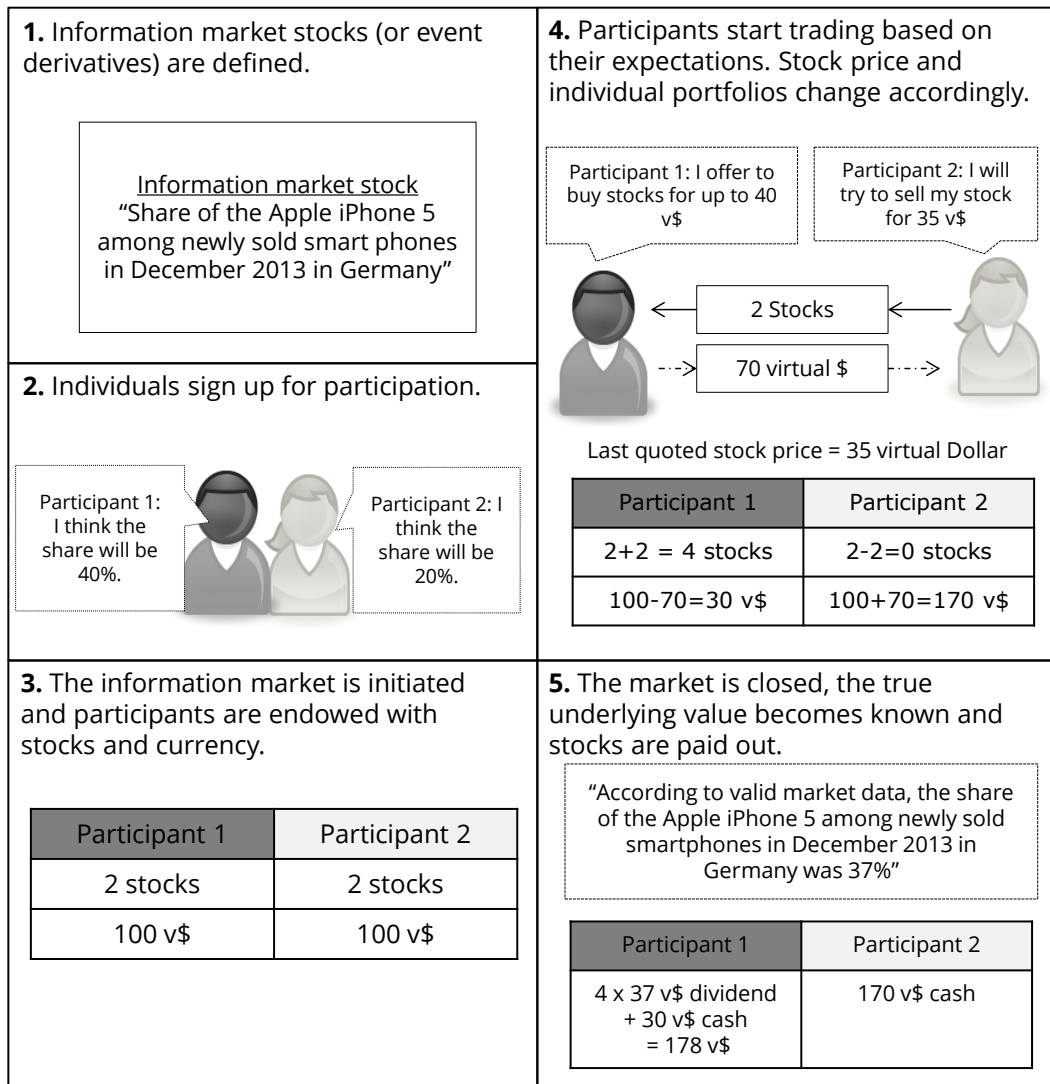


Figure 3.1.: Information market process example, from event definition to stock payout
 (Source: Own depiction)

3. Information Markets for Innovation Evaluation

result. Based on the given trade and assuming that no further trading will occur, the market prediction for the share of iPhone $\tilde{5}$ among newly sold smart phones in December 2013 in Germany would remain at 35%.

5. After December 2013, the true underlying value becomes known and stocks can be paid out accordingly. Assuming a true market share of 37%, we can observe that Participant 1 yields a higher overall portfolio value. His expectations are closer to the true underlying value, which allows him to generate larger profits than Participant 2. Assuming a different outcome in which Apple had withdrawn the iPhone $\tilde{5}$ from the German market altogether before December 2013, the resulting share among newly sold devices would be 0%. In this case, Participant 2's expectations would be closer to the true underlying value and she would accordingly generate higher trading profits than Participant 1 (170 virtual \$ vs. 30 virtual \$).

This short example illustrates that the revelation of valid expectations via trading benefits the revealer through higher trading profits. At the same time, the market's initiator benefits by gaining more accurate predictions, thus showing the information market's incentive compatibility. It also reveals that the quality of the information market outcomes depends mainly on the expectations and actions of its participants.

The efficacy of information markets builds upon the famous proposition by Nobel Prize winner Friedrich von Hayek that **markets are a relatively effective instrument** (compared to central planning) **to pass on essential information** (Hayek 1945). The efficient market hypothesis posits that prices on markets efficiently and fully reflect all available information held by market participants via the markets' trading mechanisms (Fama 1970). On a micro-level, deviations between market prices and individual expectations will prompt participants to reveal their expectations via trading. Participants will try to buy or sell information market stocks as long as they believe that current market prices do not reflect the future state of the underlying event. Empirical data from laboratory experiments and financial markets support the notion that prices in asset markets can efficiently reflect all available information (Plott and Sunder 1982; Plott and Sunder 1988; Fama and French 1992).

Historically, information markets have been used in the U.S. for betting on voting outcomes in presidential elections as early as 1868. These early markets focused on the promise of individual gains over the aspect of information aggregation, but their price quotes were still widely used by newspapers and campaigners as sources of valid forecasts.

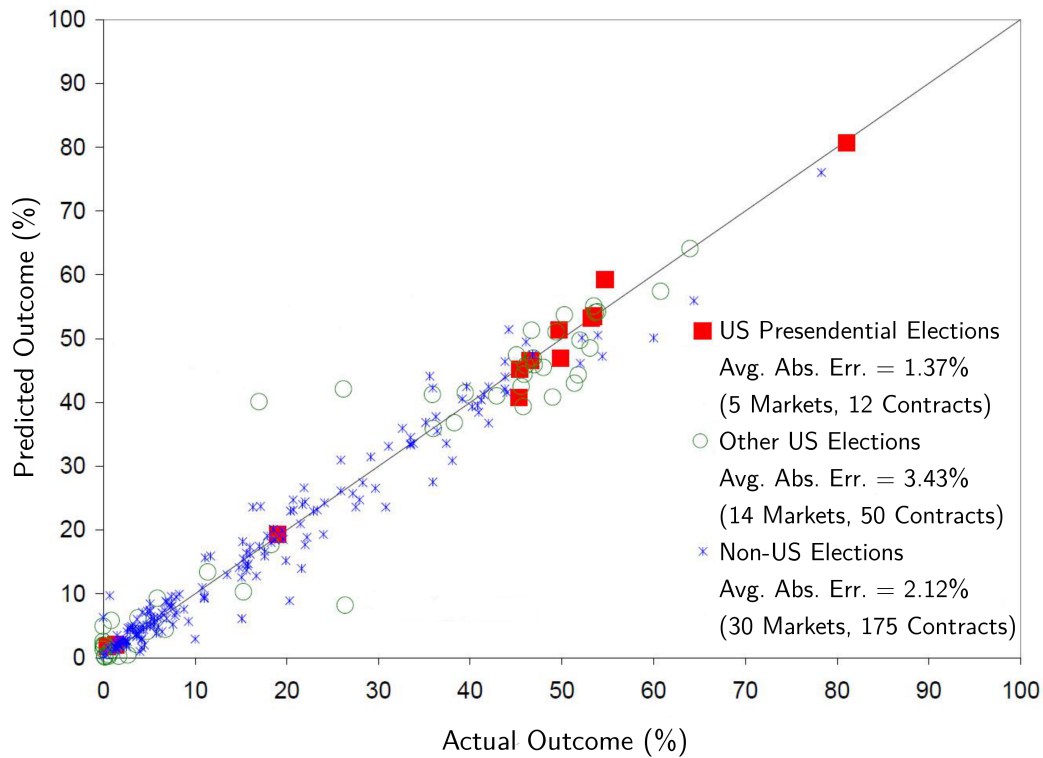


Figure 3.2.: Prediction quality in global political information markets between 1988 and 2000 (Source: Berg et al. (2008))

They reached their high point in 1916, when spending on electoral betting markets was twice the total spending on the election campaigns, and before legal restrictions forced them to cease (Rhode and Strumpf 2004).

Current scientific publicity on information markets was triggered when political markets were revived for academic purposes at the University of Iowa in 1988 (Forsythe et al. 1992). In fact, the Iowa Electronics Market has significantly outperformed major U.S. polls in all presidential elections since the market's inception. Figure 3.2 displays the absolute accuracy of 237 political information markets at midnight on the evening before the actual election. We observe considerable accuracy, as almost all predictions are very closely lined up on the 45-degree line (Berg et al. 2008).

Further results demonstrate the superiority of information markets over opinion polls for predicting election outcomes. In the same article, Berg et al. (2008) compare the accuracy of 964 opinion polls with the results of the corresponding information markets. **Information markets significantly outperformed opinion polls in the large**

majority of pair-wise comparisons. The information markets provided significantly more accurate predictions at all time points leading up to the elections.

After their success in the political domain, information markets were identified by management scholars as promising tools for business evaluation and forecasting. In 1999, the Hollywood Stock Exchange (HSX) was founded as a pioneering public information market to forecast new movie success (Keiser and Burns 1998). Movie success has historically been considered very difficult to predict. Movies are complex and gestalt-like innovative products, which inhibits valid prediction by stochastic methods (Vany and Walls 1999). Information markets have been able to absorb new movies' inherent complexity. The expectations of market participants about a movie's future success are derived from its gestalt (Karniouchina 2011). Participants include potential moviegoers and film critics, and hence incorporate ultimately sound information sources. The HSX has been very successful at predicting box office success since its introduction, significantly outperforming alternatively established methods like expert panels or opinion polls (Spann and Skiera 2003b). The authors emphasize that in contrast to expert panels like *Box Office Mojo*, information markets provide access to a larger pool of potentially valid information sources, compatible incentives for valid self-selection of participants, and the revelation of participants' true beliefs. Compared to traditional opinion polls among moviegoers, information markets foster learning from other participants via the trading mechanism and emphasize information from better-informed participants, since these are more likely to profit from the market and continue to participate (Spann and Skiera 2003a).

Ortner pioneered business applications when he applied information markets to evaluate project durations at Siemens (Ortner 1998). More corporate applications eventually followed. Information markets outperformed quantitative forecasting methods for existing product lines at HP (Chen and Plott 2002), and also outperformed Delphi methods for predicting the impact of technological change at a major German telecommunications provider (Spann and Skiera 2003a). Large technology companies such as Google, IBM, Microsoft or BestBuy have applied information markets to gather intelligence for demand forecasting and product development processes (Cowgill et al. 2008; Lacombe et al. 2007; Dye 2008).

3.2. Applications of information markets for innovation evaluation

In practice, most applications of information markets are related to the domain of innovation evaluation. Information markets for predicting election outcomes aim to capture voters' preferences for political parties or presidential candidates on election day and new movies are by their very definition innovative (though one may argue about how incrementally innovative another sequel of the "Pirates of the Caribbean"-franchise really is). However, these applications have been used to predict the outcomes of events that were very close to entering or had already entered the market. Only over the last decade have researchers begun to systematically analyze and explore the feasibility of information markets in evaluation at all phases of the innovation process. Interestingly, the earlier the phase in the innovation process, the later information markets have been applied to and investigated for use in that phase. The following paragraphs will briefly capture the current state of research on the application of information markets to innovation-related issues. We will specifically focus on illustrating the underlying prediction objects, participants, trading mechanisms and incentives for participation in innovation-related information markets. Current research based on existing applications is presented along the corresponding phase of the innovation process (see chapter 2.2).

Idea generation and assessment

Information markets for idea evaluation were first explored at General Electric (GE) (Lacomb et al. 2007; Spears et al. 2009). Until 2009, GE had run 10 information markets to source and evaluate business and technology ideas.

In the first "**GE imagination market**", the company ran an information market out of their Computing and Decisioning Sciences Technology Center (Lacomb et al. 2007). The idea market was aimed at evaluating new business ideas according to criteria that demanded advanced technology, high and reasonably fast economic impact, and close relation the organization's technology focus. Based on these criteria, five ideas were initially seeded to the information market as stocks by the market administrators. Furthermore, participants could later submit additional ideas that were first screened for similarity with existing idea stocks. The stocks were all given an initial starting price of 50 currency units and could reach a minimum value of 1 and a maximum value of 99 currency units.

The company recruited all staff from the above-mentioned center to participate in the

3. *Information Markets for Innovation Evaluation*

information market, allowing no additional external participants. Participants received an initial allowance of 10,000 units virtual market currency and a weekly allowance. The company utilized an information market software that employed a similar double-auction mechanism to what was described in the introductory information market example.

Stocks were ultimately paid out based of the volume-weighted average trading price over the final five days of market trading. The ending date of the market was not communicated to market participants in order to mitigate the danger of manipulation in the final moments of trading. Incentives were given both for submitting ideas (50,000 US-\$ in research funding to pursue the idea) and to the top traders (Apple iPods were given to the top two traders and 25\$ gift certificates to the next ten top traders).

In total, 62 ideas were submitted successfully to the market. Of 85 active traders, 24 traded at least once a day. Random gift voucher drawings were introduced for active traders during the second week of trading to increase trading activity.

After the market had finished, the outcome was compared to leadership team evaluations. The researchers found that the idea rankings from volume-weighted final stock prices were positively correlated with senior management's evaluations. A chi-square test revealed no difference in the distributions of rankings between the two evaluation mechanisms. However, the leadership team asserted that the imagination market significantly improved the quantity and quality of ideas compared to previously applied mechanisms for sourcing and evaluating new product ideas.

The same researchers provided a second study on a later imagination market at GE's nuclear energy department (Spears et al. 2009). While the basic setup of the market remained the same, some minor changes were made and the researchers' focus was place upon trader behavior during the course of the market.

Short-selling was removed because it apparently confused most traders who were unfamiliar with the underlying mechanism and thus did not understand how to benefit from short-selling. This time, a cross-functional team reviewed the new ideas submitted by market participants to evaluate their fit with the initiators' goal of finding ideas that customers would value, that would produce the best return on investment, and that should be included for funding next year.

After the market, the researchers focused their analyses on participants' trading behaviors. They found that subjects rarely traded multiple ideas that targeted the same product line but rather focused trading on one idea stock within each product line. Furthermore, idea creators bought stocks of their own ideas at significant price premiums and were less likely to sell these stocks. The researchers argue that this behavior was

fueled by a mixture of optimism, wishful thinking and the rational incentive to get their idea funded if it reached a top position after the markets had closed.

A **recent contribution** to the research on information markets for evaluating new product ideas comes from Soukhoroukova et al. (2012). The authors ran a field study at a large German industrial company. Similar to the application by Lacombe et al. (2007), their application was aimed at generating business ideas for the whole company that would bring a significant economic benefit. Furthermore, the initiators aimed to predict promising technology fields and product ideas for a specific department.

However, in their application, the authors invited participation from all company employees, ranging from white-collar employees invited via e-mail to blue-collar workers who were invited via leaflets handed out in the factories. Ideas were submitted by participants via an initial-public-offering mechanism. After submissions, other traders could buy stocks of these ideas at a fixed price per stock. If a certain threshold of stocks for a newly submitted idea was purchased within seven days, it was initially offered on the information market at the previously fixed price. The more participants actively entered the market, the higher the threshold was set for ideas to successfully enter the market. In total, 397 participants from 16 countries actively traded 100 idea stocks that successfully entered the market. However, departing from the applications of Lacombe et al. (2007) and Spears et al. (2009), payout was not determined inherently but by external evaluation of the submitted ideas. A committee of senior managers, venture capitalists and external technology experts ultimately determined an idea's value, which then determined the idea stock's payout. While expert opinions were positively related to the market's assessment of the specific new products and the broader new business ideas, market evaluations were not significantly related to the expert evaluations of important future technology fields. Still, the researchers found that participants and senior managers agreed that the idea market was a useful instrument for idea generation and evaluation and that it should be used again in the future to produce significantly more innovative ideas than the competition.

Concept evaluation and planning

At the beginning, the uses of information markets for concept testing aimed at evaluating preferences for concepts that already possessed well-defined physical and functional properties.

In their "Securities Trading of Concepts" (STOC) experiments run between 2000 and 2009, Dahan et al. (2011) focused on the **relative quality of using information mar-**

kets in measuring preferences for new product concepts, as compared to existing methods such as surveys, conjoint analysis or virtual concept tests.

The researchers analyzed the results of 11 experimental information markets for concept evaluations. The prediction objects covered a range of concepts, including bicycle pumps, messenger bags, crossover vehicles and video games. Participants were recruited from graduate and executive management courses and numbers ranged from 16 to 62 participants across the markets. All experimental markets ran for less than 60 minutes to elicit participants' preferences. In each market, continuous double-auction mechanisms were used to facilitate trading. Eight of the experiments used final market prices and three experiments used volume-weighted average prices for paying out traders' portfolios.

In the study, Dahan et al. (2011) explore three dimensions of beneficial characteristics when using information markets for concept evaluation. First, they find that **information markets are more engaging and cost efficient than alternative methods such as surveys**. Second, they show that information market results correlate highly with the results of alternative methods like conjoint analysis and virtual concept tests, but similarly fail to reliably predict future market shares over a longer period of time. Third, their experiments show that information-market prices more closely resemble the traders' expectations of others' preferences than own preferences, but that participants' trading is nonetheless biased by their own preferences.

Soukhoroukova (2006) explored the feasibility of information markets as instruments for evaluating the potential product success of new mp3-player concepts prior to market introduction.

In these experiments, the information markets required 8-12 traders while the conjoint analysis drew from 307 responses. Trading was facilitated via a double-auction mechanism and participants traded for up to 30 minutes. Portfolios were paid out according to final market prices and the best traders won gift certificates.

Soukhoroukova (2006) compared the results from the information markets with those from the well-established method of conjoint analysis and found that both approaches yielded valid and similar responses. She also found that recruiting participants for a stock market game appeared to be easier and cheaper than traditional market research methods. Participants valued the competition via trading more than responding non-competitively via a survey.

Dahan et al. (2010) further studied the feasibility of information markets for concept testing by building upon the initial STOC approach. Their study explored how **com-**

plex products with multiple features and different feature attributes can be evaluated via information markets. The researchers aimed to limit participants' exposure to the hundreds of potential features in new smartphone concepts. Instead, they tested whether prediction quality remains high when sub-groups of participants only traded sub-sets of relevant product-feature characteristics. For example, one group only traded stocks that represented different colors and memory sizes in phones, while another group only traded stocks that represented different screen types and camera resolutions.

The authors ran two separate experimental information markets, one with graduate management students and the other with senior marketing executives at a large mobile services provider. Each information market had roughly 100 participants but traders were allocated to six sub-markets with 5 to 21 traders each. Instead of trading 58 mutually exclusive new product attributes in one large market, each sub-market featured between 15 and 21 product attributes to reduce complexity for participants. Traders posted buy and sell offers via a double-auction mechanism and markets ran up to 50 minutes. Portfolios were paid out based on volume-weighted average prices and the best traders got the chance to win gift vouchers.

The researchers conclude that splitting features and attributes over different information markets can provide necessary scalability when evaluating complex new products having potentially long feature lists without reducing the viability of applying information markets, as compared to traditional methods for preference elicitation such as conjoint analysis or surveys.

Spann et al. (2009) explored another valuable aspect of information markets during the concept development and testing phase by focusing on traders as research objects. The authors started with two hypotheses: first, individuals who hold and trade on superior information about the underlying prediction target should perform relatively well in prediction markets; and second, companies will benefit from identifying and including well-informed individuals in their development evaluation processes for innovative products.

To test these hypotheses, the authors ran an information market that predicted the future success of potential movie releases. Participants were recruited via university advertisements and web-site banners on a popular German movie site. Similar to the previous studies described, the researchers applied a double-auction mechanism for trading. The markets ran for over two months and the best traders were given gift certificates. By running post-study surveys with the participants, the researchers found that more

knowledgeable and innovative individuals were significantly more likely to be among the best traders. Hence, they concluded that information markets can additionally be used as an efficient tool for identifying relevant information sources in the context of innovation evaluation compared to complex survey-based methods (Herstatt and Hippel 1992).

Market introduction and diffusion

The earliest research on innovation-related information markets targeted market success predictions. When it first began in 1999, HSX predicted the success of movies where the casts and stories were already well defined. Only later did HSX incorporate less well-defined movie elements such as the potential box-office revenues of yet-to-be announced sequels or the value of actors as individual prediction objects. HP used information markets to predict the market success of already launched products such as printing systems. The information markets outperformed official forecasts in 75% of the events studied (Chen and Plott 2002). And as mentioned before, a major German telecommunication provider forecasted the demand for new mobile data packages and services via information markets among senior management personnel (Spann and Skiera 2003a).

In short, information markets have been successfully applied for the purpose of information sourcing and evaluation at all steps of the innovation process. The table in Figure 3.3 summarizes the studies discussed in this section. In most cases, the markets performed at least on par with alternative evaluation mechanisms, and on a satisfactory level for initiators when applied in corporate environments, compared to existing methods for forecasting new product success.

3. Information Markets for Innovation Evaluation

Author	Year	Innovation-process phase	Prediction objects	Participants	Trading mechanism	Incentives
LaComb et al.	2007	Idea generation & Assessment	New business ideas for the Computing and Decisioning Sciences Department	The department's managers, project leaders, staff, contractors, interns, 150 invitees and 85 traders	Double auction, short selling allowed, dividend derived from market price	Play money, tournament with physical prizes for top three traders
LaComb et al.	2009	Idea generation & Assessment	New business ideas for the GE Energy business	Salaried employees of GE Energy, 1236 invitees and 110 active traders	Double auction, short selling not allowed, dividend derived from market price	Play money, tournament with physical prizes for top three traders
Soukhoroukova et al.	2012	Idea generation & Assessment	New technologies for the company in ten years, new product ideas for a specific product category, innovative product and business ideas for the company	All regular employees of the company (<2 bn. \$ revenue) were invited, 397 participated as traders	Double auction, IPO to submit new ideas, dividend derived from expert evaluation	Play money, tournament with cash prizes for the top ten traders
Dahan et al.	2011	Concept development & planning	Eleven concepts for new bicycle pumps	44 MBA students from MIT in two experimental sessions	Double auction, dividend derived from stock closing price	Play money, tournament with three 30\$-50\$ vouchers for the top three traders
Soukhoroukova	2006	Concept development & planning	Four concepts for new mp3 players	36 participants in four experimental sessions, two sessions with students	Double auction, short selling not allowed, dividend derived from stock closing price after an hour	Play money, tournament with dvd and cinema vouchers for best traders
Dahan et al.	2010	Concept development & planning	Six smartphone concepts with 58 mutually exclusive attributes	113 students in an experimental session and 63 professionals from a large telecommunication company	Double auction, short selling not allowed, dividend derived from stock closing price after 50 min.	Play money, tournament with 50\$ rewards for the best traders in each experimental group
Spann et. al.	2009	Concept development & planning	Number of movie visitors for future movie releases	349 visitors to a leading German movie theater website who were recruited via pop-up banners	Double auction, short selling not allowed, dividend derived from true value after two months	Play money, tournament with annual movie tickets for the best traders in each market round
Cowgill et. al.	2010	Concept development & planning	157 objects: number of future Gmail users, quality rating of future products, competitor behavior, relative usage of future features	1463 active traders comprising mainly google employees and some contractors and vendors	Double auction, short selling not allowed, dividend derived from true value after each quarter	Play money, tournament with a 10,000\$ price budget per quarter over two years running time
Chen & Plott	2002	Production, market introduction & diffusion	Sales forecasts for twelve new products; four products each predicted separately within three corresponding divisions	25-30 participants per division and 5 participants from the R&D team that developed the market	Double auction, short selling not allowed, dividend derived from true underlying value after forecasting period	Real-money prediction market where the money was endowed by the organization and paid out after the markets had finished
Spann & Skiera	2003	Production, market introduction & diffusion	Forecasts of moviegoers for upcoming movies; forecasts for usage of new mobile services at a large telecommunications company	Data from trades by over 725,000 registered traders for the movie predictions, 20 employees at the telecommunications company	Double auction, short selling not allowed, dividend derived from true underlying value after forecasting period	Play money, tournament with recognition by official ranking for the best traders in the movie market.

Figure 3.3.: Previous application of information markets for innovation evaluation

3.3. Designing information markets for innovation evaluation

The previous section illustrated that information markets have been successfully applied as information sourcing, evaluation, and expert-identification instruments at all phases of the innovation process. The examples demonstrate the potential of information markets for reducing uncertainty in the innovation processes and increasing the likelihood of identifying successful decision options during an innovation endeavor. The following sections will present and summarize the current state of research regarding the design of information markets for innovation evaluation. We will introduce the relevant design variables and discuss design choices based on prior findings in the domain of information markets.

Information markets integrate four critical design elements: participants, stocks, trading mechanisms and incentives. These four elements must efficiently support the market's underlying goals of sourcing meaningful information and providing superior evaluations, while at the same time considering organizational and environmental constraints, which may limit the implementation to certain design options (Soukhoroukova 2007).

3.3.1. Prediction objects

The underlying prediction targets are expressed via derivatives, whose value is subject to the true outcome of the underlying prediction. They are most commonly referred to as information market stock. The first critical design task when rendering the information market stocks is to decide on the form of the prediction object. Spann (2002, p. 57) differentiates between three basic types of underlying prediction values that can be transferred to innovation-related information markets, as indicated below:

1. The prediction of absolute figures, e.g. an innovation's potential customers, users, revenue, or profit in a specific time frame or the days until the completion, or the total money spent on an innovation project.
2. The prediction of relative figures, e.g. the market share of an innovation in a specific time frame, or expected preference shares for certain attributes when evaluating new product or service concepts.
3. The prediction of probabilities for binary or ordinal partitions of event outcomes, e.g. the probability that a certain sales figure will be surpassed with an innovative

3. Information Markets for Innovation Evaluation

product or that an innovation will be marketed by a certain date.

Having selected appropriate prediction objects, the second critical task is to create non-ambiguous prediction targets. The underlying prediction target must be defined in such a way that future events are identically perceived by all market participants and the market initiator. For example, predicting the market share of “the new BMW” may be perceived ambiguously if more than one new model is being introduced to the market. Similarly, the geographical and temporal context of the prediction needs to be well defined. In the same vein, relational conditions must be clearly carved out when predicting percentage shares, e.g. “The market share of all new BMW 3-class models during the first half of 2013 in Germany among all privately registered new vehicles.”

Arguably the greatest challenge when designing information market stocks for innovation evaluation dealing with **the lack of ex-post validation for early phase predictions**. When predicting the market success of innovations that have already been introduced to the market, final stock prices can easily be linked to **true market figures** after the predictions have taken place. However, when evaluating innovations at earlier phases of the innovation process, true values to determine stock dividends are much harder to define and measure. Several theoretical and practical problems arise when designing an information market to evaluate new product ideas or the potential of alternative concepts. How can an information market’s initiator truly measure the business potential of a new product idea, which should determine the related stock’s ultimate payout? First, it will likely be difficult to find an unquestionable measure for “business potential.” Second, even if such measure is identified, it may not materialize for quite a while, as innovation development from the initial idea to final product usually takes many years. Finally, it will likely be hard to discriminate the extent to which a specific idea contributed to an innovation that can ultimately enter the market. Prior applications of information markets for idea and concept evaluations have approached the issue by using market-inherent determinants to pay-out stock prices.

For example, Lacombe et al. (2007), Spears et al. (2009), and Dahan et al. (2011) used volume-weighted average stock prices of the final market trades prior to closing to determine their final value and pay out participants’ portfolios. Similar to these applications, Dahan et al. (2010) stopped their markets to evaluate product attributes randomly after 30 to 45 minutes run time without previously notifying the participants and then paid out portfolios according to the last quoted stock prices.

Recent experiments have compared the predictive quality of information markets when

3. *Information Markets for Innovation Evaluation*

paying out portfolios according to the abovementioned market-inherent stock valuations, to paying out portfolios according to their true underlying value. The results demonstrate that determining payouts by volume-weighted average prices and final market prices can yield similarly good predictions compared to paying out stocks based on true outcomes (Slamka and Jank 2009).

However, Ottaviani (2009) rightfully claims that such payout schemes are textbook examples of “Keynesian beauty contests,” where participants do not predict fundamental values but rather what other participants believe the fundamental values to be. Instead, he proposes using proxies of true outcomes to determine stock payouts, based on some ex-post performance measure. For example, initiators could resort to expert committee evaluations or anonymously run parallel markets with distinct and mutually unknown participants where final market prices in one market determine the stock payouts in the other (Ottaviani 2009).

Still, one may critically mention that such proposals move predictions of others’ expectations outside the market rather than abolishing them. Soukhoroukova et al. (2012) paid out idea stocks in an industrial application based on senior management evaluations. They mention that correlations between market prices and external evaluations were low on average, but that their scheme was still widely accepted by market participants. Finally, innovation management research frequently relies on expert or senior management evaluations as validation criteria for innovation value in the absence of a true, underlying value, which supports the idea applying expert evaluations as proxies in cases where true outcomes can not be measured (Kristensson et al. 2004; Franke et al. 2006; Poetz and Schreier 2012).

We conclude that the biggest challenge in designing information market stocks and defining their fundamental value arises when evaluating innovations at the idea or concept stage. First, designing the stock in a way that it can be commonly understood is likely to be more difficult in the early stages, where not all characteristics of the idea can be unambiguously described. This is particularly true for information markets that recruit traders from different organizational domains possessing very heterogeneous knowledge. Second, finding external validation criteria to determine the stocks’ true values after trading is likely impossible for stocks that derive value from the market potential of ideas and concepts. Yet, novel research has shown that market-inherent payout schemes and external proxies such as expert committees can provide meaningful alternatives for determining the stocks’ fundamental values.

3.3.2. Participants

Information market participants are a key element to information market success. First, they are the primary source of the information that is aggregated in the market prices. While other sources of information are often publicly provided by the information market initiator via in-market news windows or real-time tickers (Spann 2002, p. 197), trading is essentially triggered by discrepancies in participants' private expectations (Luckner 2008, p. 20). As discussed earlier, participants engage in trading because they assume that they possess information that allows them to harvest trading profits.

In the context of innovation evaluation, traders may be recruited from all of the domains that were discussed in Section 2.3.1. Including outside views from customers, suppliers and experts is similarly associated with superior performance (Dye 2008). Although participants' information remains sticky and tacit, the market mechanism can act as a catalyst for revealing and aggregating it. Yet, different types and compositions of participants may be appropriate for different stages of development.

Initiators need to be aware that information aggregates and **final outcomes are, by definition, transparent among market participants**. All traders may use these results alike, which may cause the flow of knowledge to competitors, suppliers and customers if the participants have been recruited from these groups. All things being equal, the participation of these groups may result in increased competition or lower bargaining power, e.g. if competitors obtain higher marginal utility from the market signals than the initiating organization. Especially in the phase of idea generation and evaluation, some participants may have the incentive to exploit the proposed ideas, which could negatively impact the initiators' exploitation goals. For example, in the markets of Lacombe et al. (2007) and Spears et al. (2009), employees evaluated ideas that were submitted by themselves and the highest ranked idea received 50,000\$ in research funding. Such a setup creates incentives for idea owners to trade according to what is best for them personally, rather than what would have the most positive business impact for the organization.

Such agency problems may be less relevant in applications during later phases of the innovation process, when the innovation is more mature and closer to market introduction. Many relevant assets for innovation success may have already been acquired and protected, such as access to physical resources, technologies, trademarks or distribution channels.

Furthermore, and especially so during early phases of evaluation, **it may be par-**

ticularly difficult to recruit traders who actually possess valid information regarding the broader and future market relevance of the proposed ideas. We illustrated in the second chapter that this is not a problem that is exclusive to information markets as a method for evaluating ideas. Still, initiators must consider that individuals may be more likely to reveal current personal preferences than valid expectations of broader market needs. Dahan et al. (2011) show that participants are indeed systematically biased towards their own preferences when trading in information markets. In another study, however, the researcher points out that information markets ultimately achieve reduction of personal preference biases compared to aggregating pre-market beliefs (Dahan et al. 2010).

Besides having participants who provide valuable information from multiple domains, it is also important to have a sufficient amount of traders to provide the market with liquidity. If all participants knew the stock's true underlying value perfectly, none of the traders would have an incentive to buy or sell stocks, rendering the information market futile. In fact, many researchers have stressed the advantage of including a significant number of uninformed traders in the market because that would enable traders with highly diagnostic information to better profit from revealing their information (Wolfers and Zitzewitz 2004).

Indeed, Lacombe et al. (2007) specifically included traders who were not expected to contribute meaningful information in order to provide market liquidity. Informed traders will more likely act upon their superior information and *make markets* actively, while uninformed traders are more likely to engage in trading via price taking, as Oliven and Rietz (2004) show in their analysis of trader behavior in information markets.

Finally, the **fear of having participants who actively manipulate information market outcomes** has previously hindered the adoption of information markets for related applications such as foreign-policy forecasting. Critics have proposed that manipulative traders could aim for and cause systematically skewed estimations (Pearlstein 2003).

So far, theoretical insight into the impact of manipulation on prediction accuracy is scarce. Robin Hanson and colleagues provide two experiments demonstrating that (1) manipulators do not significantly deteriorate market accuracy if the traders can identify their presence (Hanson et al. 2006), and (2) external market observers will draw equally efficient inferences from market prices if manipulators are present (Hanson et al. 2007).

However, in another experiment where manipulation incentives were significantly higher, market prediction quality deteriorated in cases where true beliefs pointed to high stock prices and manipulators engaged in trading to lower stock prices (Deck et al. 2013). Taking these findings into account, information markets appear quite robust in the presence of manipulative traders. However, results may deteriorate if incentives for truthful revelation are overruled by higher incentives for manipulation for a sufficient number of participants.

3.3.3. Trading mechanisms

Information markets only aggregate and reveal dispersed and heterogeneous private information via their price mechanism if market participants successfully engage in transactions based on their expectations. The trading mechanism is therefore another integral part of an information market. Most importantly, the trading mechanism controls the speed and degree of independence by which market participants can execute transactions at the desired quantities or prices. The mechanism consequently influences the liquidity of the traders' assets and impacts the speed and magnitude by which traders actions can influence stock prices. The majority of prior applications in laboratory experiments and field applications have employed one of two main classes of market mechanisms. Trading has either been facilitated via continuous double-auctions (CDAs), which allow participants to exchange information market stocks and currency, or has used incorporated automated market makers, which automatically provide liquidity (trading opportunities) whenever participants decide to buy or sell information market stocks.

CDAs have been coined as the standard mechanism for information aggregation via markets (Ledyard et al. 2009; Healy et al. 2010) due to their long history in experimental markets and information markets (e.g. Smith (1962), Forsythe et al. (1992)). Accordingly, CDAs were used as trading mechanisms in all cases of information market application for innovation evaluation introduced in the previous section. As mentioned earlier, CDAs facilitate trading via an open order book. Traders post buy and sell offers in the order book. When prices of new buy offers are equal to or higher than prices of outstanding sell offers, or if prices of new sell offers are equal to or lower than outstanding buy offers, trades are executed between the parties involved.

Automated market-maker mechanisms facilitate traders' actions via an artificial trading agent that aims to continuously provide a sufficient number of buy and sell offers

3. *Information Markets for Innovation Evaluation*

close to current market prices, and which market participants can then interact with. If market participants react to the market maker's buy offers, the asset price will rise and if market participants react to the market maker's sell offers, the asset price will fall. In any case, the market maker will update its current offers after asset price adjustments to maintain the market's liquidity. Market makers' central benefit lies in the continuous provision of liquidity. In a thin market in which few participants engage in trading, participants can easily reveal information via a market-maker mechanism because it never requires a another participant as counter party in order to carry out a successful trade. These trades would be less likely to be executed in a similar thin CDA-based market (Hanson 2003). While market makers are considerably easier than CDAs for traders to understand, they are subject to more endogenous and complex rule sets (Klingert 2013). Automated market makers rely on algorithms that automatically adapt the stock prices in case of participants buy or sell stocks.

There is a lively debate among experimental market researchers regarding the benefits of the different trading mechanisms and their relationships with markets' environmental characteristics. Healy et al. (2010) conducted an extensive experiment comparing the prediction errors of different market mechanisms with alternative aggregation methods such as surveys. They found that automated market makers significantly outperform CDAs in complex market environments in which outcomes may take multiple states. Based on their results, the authors argue that (1) CDAs take two participants for a successful trade, which is more labor intensive and may reduce information revelation and (2) CDAs are more susceptible to far-off last reports because any single trade can produce large shifts in stock prices, which is naturally prevented when using automated market makers.

Although current research on information market applications for innovation evaluation mainly featured CDAs, it would be wrong to conclude that CDAs are the trading mechanism of choice in this context. First, the mechanism may have been overly represented simply because most software providers have only started to introduce automated market makers as trading mechanisms over the last couple of years. Today, however, most professional providers of information market software such as Inkling or Crowdworx rely on some form of automated market maker to facilitate trading. According to a senior manager at Crowdworx, clients particularly appreciate that information markets provide meaningful and fast aggregation with very few participants and are quite robust against very quick price changes. Especially for early phase innovation evaluations, Crowdworx

3. Information Markets for Innovation Evaluation

clients often invite only a limited set of in-house participants to better protect intellectual property.

Finally, financial markets often provide the opportunity to post offers or commit trades by borrowing money and stocks. **Credit-based trading** has been present in the domain of information markets since they were first applied (Forsythe et al. 1999). From the previously discussed applications, only Lacombe et al. (2007) allowed short-selling of stock and none of the applications allowed participants to have negative cash accounts. Short-selling was disabled mainly because the concept is difficult to understand for novices to financial markets (Chen and Plott 2002; Spears et al. 2009).

On the one hand, buying or selling via lending money or stocks may benefit market efficiency by allowing traders to reveal information without the necessity of having the corresponding capital. On the other hand, allowing credit-based trading may motivate traders who have previously lost money and stocks to resort to borrowing for trading, as was discussed with Prospect Theory (Miller and Chen 2004). Extensive credit allowances may also have detrimental effects on market efficiency, if building on the basic assumption that the traders previously lost money because they provided information that did not improve market predictions.

One reason for not allowing credit-based trading is the lack of experience in using it. The information market applications discussed earlier were mainly first-time applications in their respective organizations and initiators may have been shy to test novel extensions such as automated market makers or credit-based trading. Another reason could be that allowing credit creates a problem in virtual currency markets. How should initiators act upon negative virtual currency accounts after stocks have been paid out? If credit-based trading is disabled, traders may, at worst, end up with a portfolio value of 0, in the case that they spend all of their cash on stocks that did not pay out any dividends. When allowing credit-based trading, however, traders could end up with negative portfolios after trading, and initiators would likely be reluctant to penalize negative portfolios after the information markets have finished.

In sum, credit-based trading could fuel information revelation but could also induce the wrong subjects to overly engage in trading. Additionally, it could negatively impact the current perception of information markets for corporate innovation evaluation as a playful and enjoyable instrument. It seems to have been a sensible choice for the cited organizations to restrain themselves from implementing it.

3.3.4. Incentives

Information market incentives govern the likelihood that participants will join the market and reveal truthful information via trading. Initiators should aim to align incentives for joining with those for truthful revelation. However, high participation rates could have negative effects on information market quality if additional traders do not help to improve prediction quality. The information market applications that were previously discussed did not offer specific incentives for signing up to participate but offered incentives that were only conditional upon good trading performance.

In general, incentives for the revelation of truthful information are tied to final portfolio values after the underlying outcomes are quantified and the stocks are paid out accordingly.

At first, the initiator must decide whether participants will use their own currency in the information market, or if they will be exempted from any direct financial risk. Early political stock markets required participants to invest private funds (Forsythe et al. 1992), but limited maximum investment to no more than 500 US-\$.

During the last decade, **two studies have explored whether real-money markets outperform play-money markets** that do not require private investments. Servan-Schreiber et al. (2004) compared two popular online information markets for sports forecasting, one based on real money and the other on play money. They found that the predictions in both markets were equally precise and significantly outperformed individual expert predictions. They conclude that information markets do not necessarily need to provide the incentive of monetary gains as long as other incentives motivate participants to engage in trading and to outperform their peers. Another study analyzed similar data from the same information markets but from a broader domain of topics and over a different time frame (Rosenbloom and Notz 2006). The authors found that the results were equally accurate, but only for sports predictions where participation was similarly high in the play-money markets. However, the play-money market performed worse in less populated special-interest markets. The authors propose that either the retention of losers in the play-money market or the disproportionately larger number of marginal traders in the real-money markets could explain the discrepancies (Rosenbloom and Notz 2006).

Taking these findings into account, it may be important to assess what potential risks arise from introducing a real-money information market. If traders are allowed to buy any amount of shares in such a market, the participants with significantly more financial resources might exert significantly more influence on the market prices. Drawing from

3. Information Markets for Innovation Evaluation

the results of Deck et al. (2013), these traders could consequently be able to manipulate market price interpretations. Such a danger can be mitigated by limiting overall individual investment. Furthermore, requiring participants to invest real money would eventually cause some participants to lose money, in the case the market derives liquidity only from participants' investments. It appears questionable whether information markets would still be broadly perceived as a enjoyable experience by the vast majority of participants. Early-phase innovation evaluations that lack valid ex-post criteria to determine stock payout could particularly suffer from being perceived as unfair in the case that participants lose real money based on evaluations that do not necessarily reflect "true value."

After the initiator decides to base incentives on participants' real-money investments or personal funds in a play-money market, an appropriate **incentive scheme** needs to be designed. Three options are most commonly used in experimental, public or corporate information markets (Spann and Skiera 2003a; Luckner 2008, p. 81):

1. Participants' prize money or probability of winning a prize is deducted from their portfolio after the stocks have been paid out via a transformation function, e.g. 1 US-\$ per currency unit in the portfolio after the stocks have been paid out.
2. Prizes are paid out according to participants' rank order after the stocks have been paid out, e.g. 500 US-\$ to the best, 300 US-\$ to the second-best, and 200 US-\$ to the third-best trader.
3. Prizes are awarded independent of information market performance, e.g. a flat fee is paid to every participant.

In their field study, Luckner (2007) found that only rank-order and performance-compatible schemes lead to good prediction results, stressing the necessity for incentivizing truthful information revelation. Yet, the researchers also find that incentives via portfolio transformation may reduce trading activity due to risk aversion. Building on previous work showing that individuals are, on average, mildly risk averse, the authors argue that individuals will be more reluctant to engage in trading when each trade has a direct impact on unambiguous cash holdings.

In our examples of information markets for innovation evaluation, we find 9 studies that employed rank-order tournaments to distribute incentives among participants, and only one very early application by Chen and Plott (2002) that cashed out participants by transforming their final portfolio value into real currency. It appears sensible to use

3. Information Markets for Innovation Evaluation

rankings to raffle prizes where the number of entries you get is determined by your rank; particularly in markets with many participants. Initiators may find it necessary to cap incentives at a certain threshold of participants in order to limit prize money but still create attractive incentives for the best traders.

In sum, performance-related incentives are a necessary precursor for the aggregation of information by information markets. Yet, previous applications and research have shown that it may be harmful to require participants to invest private funds for trading. In most of these applications discussed, initiators handed out play money to traders and distributed incentives among the top ranking traders after the markets had finished.

4. Judgmental Biases in Innovation Evaluation

In this chapter, we will discuss the origins and impacts of important judgmental biases in the context of innovation evaluation. We will specifically focus the impact of biases on innovation-related decision making. We begin by highlighting, in Section 4.1, the increasing focus in economic research on agent behavior that departs from the traditional economic assumption of rationality. Next, in Section 4.2, we will introduce the reader to biases with specific attention to their origins in cognitive heuristics and motivation. A literature review is conducted in Section 4.3 to identify the most important biases in innovation management, their impact on innovation (evaluation) decisions and their respective origins, so as to further narrow the focus for the subsequent empirical research. Finally, in Section 4.4, we draw on current findings from information, financial, and experimental markets, and on the insights from the previous chapter, to discuss the potential impacts of judgmental biases in applying information markets for innovation evaluation.

4.1. (Ir)rationality in economic behavior

“Traditional economic theory postulates an ‘economic man’, who, in the course of being ‘economic’, is also ‘rational’. This man is assumed to have knowledge of the relevant aspects of his environment, which, if not absolutely complete, is at least impressively clear and voluminous. He is assumed also to have a well-organized and stable system of preferences, and a skill in computation that enables him to calculate, for the alternative courses of action that are available to him, which of these will permit him to reach the highest attainable point on his preference scale.” (Simon 1955)

Popular opinion from economic research has historically been guided by the assumption that most economists perceive human beings as fully informed, exclusively rational, and

4. Judgmental Biases in Innovation Evaluation

purely egoistic actors, who often strictly aim to maximize monetarily-valued utility, and that such a perception is hardly reflected by human nature (Eichner 1983). Based upon these beliefs, public critics have often mildly derided the results of economic research as largely unattached to the true mechanisms that drive economic behavior, and have hence, criticized these results for their lack of external validity (Trivers 2011).

Yet, when examining the work of well-known economists, one can find a great deal of evidence that only a few scholars believed that human subjects are truly fully-informed and rational agents. Adam Smith, in *The Theory of Moral Sentiments*, referred to human behavior as a continuous struggle between passions and impartial objectivity. Smith stressed that passions often overrule the impartial perspective, leaving it impotent (Ashraf et al. 2005). Even Jeremy Bentham and his student, John Stuart Mill, often portrayed as the inventors of the “economic man”, were well aware that their parsimonious renderings of individual utility lacked external validity because of its complex psychological underpinnings and the large variances in human character (Persky 1995; Camerer and Loewenstein 2004).

As indicated by the introductory quotation, however, most economic theory indisputably draws from the abovementioned concept of the “economic man” (Eichner 1983). Simon (1959) points out that classical economists had “...largely been preoccupied with normative macroeconomics ...”, and that they considered rationality a reasonably facile assumption for deducing theoretical models. Simon furthermore agrees with the critics and stresses that the extension of these assumptions to the study of micro-economic relationships yields especially fundamental problems because it “...neglects the processes and mechanisms through which [individual adaption to subjective environments] takes place.” Rather than building deductive theory from untested assumptions, Simon urges empirical studies on the reality and behavioral implications of subjective utility, information and environments (Simon 1955). Awarded with a Nobel Prize in 1978, Simon is often portrayed as a founder of behavioral economic research. His criticism of classical economic theory and his subsequent introduction of the “bounded rationality” concept are considered as the forefront of overcoming the gap between traditional assumptions of economic behavior and the subjective realities of economic actors (Kahneman 2003). Indeed, the second half of the last century produced a whole new stream of research on the mechanisms and processes of human decision making and their impact on economic behavior. Scholars have identified several domains in which empirical results have departed from the “economic man” assumption. Today, arguably, the most impact is attributable to results that identified human “shortcomings” in decision making under

uncertainty. Most prominently represented by the research of Daniel Kahneman and Amos Tversky, psychologists have shown that humans make systematically erroneous decisions, if one applies the expected behavior of the “economic man” as a yard stick (Sent 2004). In the tradition of Simon (1959), however, subjects who depart from the “economic man” should not be confused with irrational subjects if subjectively perceived environments differ from objective environments. Therefore, many researchers agree that such departures or biases should be studied without necessarily branding them irrational, especially as the rationality terminology remains largely ill-defined (Milijkovic 2005).

4.2. Fundamentals of bias in decision making

Human biases can be described as skewed subjective perceptions of objective realities. Biased subjects are deterred from accessing, memorizing or utilizing objectively diagnostic information, but instead yield alternated versions of *reality* after mental processing. Biases result from invalid external information cues or invalid cue processing. In the latter case, subjects transform valid external information cues into invalid internal reflections of these cues. Actions are considered biased if subjects apply invalid information to form expectations, make evaluations, or render decisions that would be considered incorrect based on an objective assessment of publicly available information. There has been an ongoing debate on the origins of biased decision making by human agents. The predominant view in behavioral economics attributes biased decisions to systematic cognitive processes that underlie human information retrieval and processing (Nisbett and Ross 1980). In social psychology, however, a considerable strand of research has argued for considering motivational factors or affective states as important drivers of biased decision making (Kunda 1990). Some researchers have long called for an increasing integration of cognitively and motivationally-induced biases (Tetlock and Levi 1982), while others add that motivational and cognitive biases have distinct origins and characteristics, and that more benefit comes from analyzing and understanding both concepts attributively and interactively (Moore and Healy 2008; Hilbert 2012). In the following sections, we will introduce both approaches to explain the origins of biases, and highlight their specific formation and characteristics.

4.2.1. Biases from heuristics and cognitive processing

Biases from heuristics and cognitive processing result from skewed information retrieval or processing. The popular *heuristics and biases* terminology was coined by Daniel Kahneman and Amos Tversky, who studied phenomena that indicated ineffective human information use, leading them to develop theoretical approaches to help understand the underlying mechanisms (Goldstein and Hogarth 1997). Cognitive biases occur when subjects systematically access and process distorted information via heuristics from memories or environmental cues (Kahneman 2003). Several such systematic distortions have been introduced to the current body of research. In the following paragraphs, a few cognitive biases will be introduced. We will also highlight the notion that heuristic decision making is not necessarily inefficient or biased.

First, humans frequently retrieve skewed information that overly contributes to the formation of specific beliefs. Subjects tend to equate the presence of information to the relevance of that information. This will frequently, for example, lead subjects to overestimate the relative probabilities of certain events that have more exposure than others due to media coverage and which lack valid data. Anchoring effects document the systematic misuse of information. Subjects make evaluations based on present cues or anchors, even if these cues are objectively unrelated to the focal evaluation task or imply implausible results (Mussweiler and Englich 2005).

For example, subjects overestimate the proportion of fatalities related to accidents compared to more subtle illnesses because accidents are more frequently visualized and sensationalized by the media (Serfas 2011). Similarly and more closely related to innovation evaluation, successful new products are more often featured in the press than flops, which may induce an overestimation of success probability for new products in general. Furthermore, anchoring may prevent innovators from discriminating between potentially successful innovations. An innovator may perceive the degree of an innovation's novelty as a positive per se because it may underline the personal creative effort. However, the innovator would likely underestimate the potential negative effects of novelty, e.g. having to explain the innovation or train potential customers.

Of course, **heuristics do not always lead people to *bad* or *flawed* decisions**. Many researchers have found empirical evidence that simple heuristics can, in fact, be very effective in human decision making (Gigerenzer and Todd 2008). For instance, Scheibehenne and Bröder (2007) show that lay people can provide similarly good predictions for the outcome of tennis tournaments compared to official experts by merely recognizing player names. In another field experiment, forecasts based on name recog-

dition were not only as accurate as statements about voting intentions in predicting federal and state elections in Germany but even worked well with very “lousy” samples (Gaissmaier and Marewski 2011). Such findings underline the assertions of prominent critics like Gerd Gigerenzer, who suggest not to focus the study of heuristics on flawed results (Gigerenzer and Brighton 2009); a focus on the biased use of heuristics does very little to explain the ecological validity of heuristics in general (Gilovich and Griffin 2002). In fact, “[...] equating limits with failure, and lack of limits with success, may underlie a deep misunderstanding about the consequences of omniscience, which may inhibit the retrieval of really relevant information]” (Gigerenzer and Todd 2008).

In sum, a large body of research has explicitly studied heuristics that produce cognitive biases, while other researchers argue that such a focus systematically misrepresents human decision-making heuristics by focusing on the below-average distribution tail of heuristic decision making (Gigerenzer and Brighton 2009). We agree with the proponents of studying cognitive biases, who point out that **certain biases arise systematically by flawed heuristics** (Kahneman 2003) and that flawed heuristics imply distortions in information acquisition and processing. It is by no means uncommon in the business world or science to focus a study on distortions in order to ultimately overcome them (Serfas 2011).

4.2.2. Biases from motivation

The previous section highlighted that subjects often make ill-informed decisions, even if they were properly motivated. In this section, we will focus on **motivation**. The following paragraphs will provide a brief introduction to the concept of motivation and show that motivation may preemptively interfere with the rendering of valid decisions because it forms uncorrelated subjective goals (Kunda 1990).

Motivation is commonly referred to as subjective goal formation. Motives are related to emotions; they reflect preference or susceptibility for specific classes of incentives of similar background. Such preference finds expression in analogous disposition to perceive and evaluate situations. The dispositional character of motives allows observation of motivations only if they are stimulated by motive-relevant situations. A motive manifests in a subject’s tendency to observe and engage in situations in a specific manner (Rothermund and Eder 2011).

Motives are considered relatively stable over time. They are inherited or learned. Figure 4.1 presents the basic model of traditional motivation theory. Motivations are con-

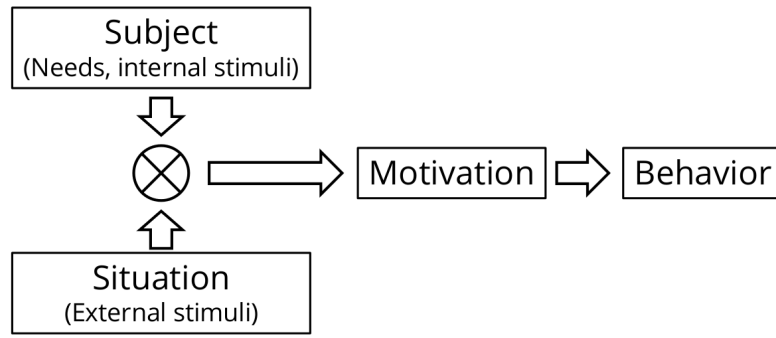


Figure 4.1.: Classical model of motivation from motivational psychology (Source: Rheinberg (1997))

sidered structures that are embedded in a person and which are triggered by relevant stimuli in a given situation. Such stimuli can be located inside or outside the subject. Intrinsic motivation draws from internal stimuli by which an action itself provides the focal subject with pleasure or satisfaction. Extrinsic motivation draws from external stimuli that drive action via material or non-material rewards. Material rewards can relate to the acquisition of financial benefits whereas non-material rewards can relate to aspects of personal security or external appreciation (Heckhausen and Heckhausen 2006). Only the interaction of subject, stimuli and situation create motivation. Consequently, the motivation controls or influences behavior, i.e. perception, evaluation and action (Rheinberg 1997).

Biases from motivation in (innovation) evaluation occur when individuals pursue goals that are independent of objectively accurate decisions or evaluations. In this case, intrinsic and extrinsic stimuli provide incentives that bring the subject away from making objectively valid evaluations.

Pyszczynski and Greenberg (1987) developed a model that describes how intrinsic stimuli for self-enhancement can influence motivation to bias decision making. According to their model, the goal of maintaining high levels of self-esteem and the human tendency towards **biased hypotheses and testing** creates biases that lead to the creation of overly positive self-images. Biased hypothesis testing refers to the assumption that decisions are generally formed based on developing a hypothesis and then collecting information to test it. However, humans will tend to search for information that supports the hypothesis because they are more likely to retrieve case-positive information and put more value on information the earlier it is retrieved Pyszczynski and Greenberg

(1987). For instance, subjects who were told they performed badly in an intelligence test attributed more validity to critical reports than supportive ones for the importance of applying intelligence tests during job interviews (Pyszczynski et al. 1985).

Shepperd et al. (2008) highlight that intrinsic stimuli often motivate self-enhancement, which may lead to biased evaluations. Subjects make self-serving attributions because they benefit in self-worth. Hence, they assume responsibility for the desired outcomes but neglect responsibility for outcomes that are not desired. In the context of innovation evaluation, such self-serving effects may prevent participants from validly assessing an innovation's value. Innovations that benefit other participants more than the focal subject are thus more likely to receive bad reviews by the subject, even if the task urges objective evaluation.

In sum, motivations are essential drivers of human behavior. Biases from motivation occur because subjects are exposed to stimuli and situations that drive motivation, and behavior that hinders the rendering of objectively valid evaluations. Alicke and Sedikides (2009) found very strong empirical evidence supporting the important role of motivation in imposing goals that prevent subjects from making objectively valid evaluations.

4.3. **Systematic literature review of research on biases in innovation management**

As discussed in the previous chapter, the effectiveness of innovation evaluation largely depends on identifying individuals who hold valuable information and who desire to apply this information to the evaluation task at hand. However, the previous section has highlighted the fact that subjects may be hindered from making valid decisions due to biased information retrieval and processing. Different biases may have distinct origins in motivation and cognitive processes, or in informational environments that deprive cognitive heuristics of their effectiveness. The following section aims to highlight the origins and manifestations of such biases in innovation management decisions. We will especially focus on decisions that are closely related to the evaluation of innovations. The goal of this section is to provide answers to four central questions:

1. Which biases are most relevant in the innovation management process in general and in innovation evaluation tasks in particular?
2. How are these biases conceptualized?

3. To what extent are these biases rooted in motivational and cognitive processes?
4. Which impacts on innovation-related decision making can be observed?

The following section will address these central questions by conducting a systematic literature review. We will present the reader with a thorough overview of the most important biases in the domain of innovation management and their respective origins in cognition and motivation.

4.3.1. Methodology

We began the literature review by selecting a set of relevant and sufficiently well-regarded scientific journals. The selection was carried out by referring to the VHB Jourqual2 rankings by the German Academic Association of university professors in business administration. The VHB Jourqual2 ranking reflects quality perceptions of national and international scientific publications from the perspective of university professors in German-speaking countries and is arguably the most popular tool in Germany for ranking journals and quantifying scientific publication quality and performance in economic disciplines (Schrader and Hennig-Thurau 2009). In order to maintain a high level of input quality, we decided to cap the journal selection at A-ranked journals, or the second highest quality level out of a total of six quality levels (A+ to E). The selection was further narrowed down by only selecting journals from the VHB-Jourqual2 sub-rankings for *Marketing*, *Technology and Innovation Management*, and *General Business Studies*, which are either the most closely related to innovation management topics or cover all areas of business administration at a sufficient quality level and are likely to provide valuable publications for further analysis in the area of innovation management (Schrader and Hennig-Thurau 2009). (We scanned the 2011 issues of all journals in this selection and focused the analysis on those journals that featured at least one article per issue that dealt with biases in innovation) After the journals were selected, we set the starting date at 1991 in order to focus on recent publications, yet control for emerging and fading trends in research during the last two decades. We then scanned the titles, keyword listings, and abstracts of all publications within the selected journals for the inclusion criteria keywords of *bias(es)*, *heuristic(s)*, *judgment(s)*, and *innovation (management)* to create a list in which each entry represented a single scientific article that could be examined more closely. For better clarity and readability, we will refer to the individual articles via numerical citations that correspond to the literature overview table in the Appendix.

4. Judgmental Biases in Innovation Evaluation

Journal title	VHB Jourqual2 rank
Academy of Management Journal (AMJ)	A+
Management Science (ManSci)	A+
Journal of Marketing (JM)	A+
Journal of Marketing Research (JMR)	A+
Strategic Management Journal (SMJ)	A
Research Policy (RP)	A
Journal of Business Venturing (JBV)	A
Journal of Product Innovation Management (JPIM)	A
Organization Science (OrgSci)	A

Table 4.1.: Journals selected for the systematic literature review

4.3.2. Results

Based on abovementioned search criteria, we identified 112 articles from the nine pre-selected journals. Afterwards, we examined these 112 articles to see whether they substantially addressed the relationship between judgmental biases and decision making in innovation evaluation. This examination resulted in a final selection of 75 articles. These articles were then coded by (1) the focal bias(es) they were addressing, (2) whether the bias(es) were rooted in cognitive and/or motivational origins, (3) methodology and (4) unit of analysis.

The allocation of biases to articles was carried out by the author and two research assistants who wrote their graduate thesis about biases in innovation development. Independently, all three reviewers wrote down biases that were referred to in the articles based on the bias definitions that will be provided below. Afterwards, the three lists were integrated by taking the intersection of all three analyses. Articles that did not provide a bias that was shared at least by two reviewers were discussed in detail to allocate the proper bias(es). The following paragraphs will show that some ambiguities exist among researchers as to how categorize the focal biases of their studies. Bias origins were coded similarly by matching their description to origins in cognition and/or motivation. The review will show that scholars most commonly argue for combinations of cognitive and motivational influencing factors behind biased decisions.

Biases in innovation management and evaluation

On average, we extracted 9.4 articles from each journal. Most articles were published in ManSci (21) and the least number of articles was found in RP (1). Between 1991 and

4. Judgmental Biases in Innovation Evaluation

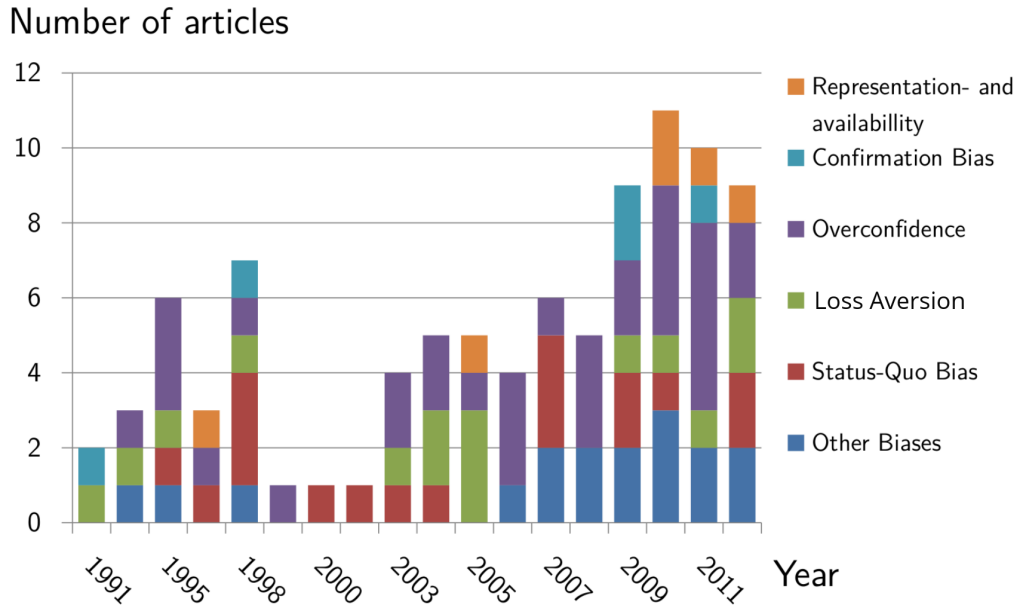


Figure 4.2.: Number of articles on biases in innovation management, by bias and year (Source: Own depiction)

2002, we see little change in the average number of articles published per year, which is 1.6. Strong growth in publication numbers starts in 2003 and rises continuously, with the average number of articles per year climbing to 5.7 between 2003 and 2012. The awarding of the Nobel Prize to Daniel Kahneman in 2002, whose research pioneered the fundamentals of decision biases and covered many facets of biased decisions' impacts in economic environments (Tversky and Kahneman 1975), may explain the steep increase in bias-related research thereafter. The first research question of the literature review addresses the distribution of distinct biases in innovation management research articles. Figure 4.2 visualizes the prevalence of the biases in all selected journals from 1991 to 2012. The graph displays the six biases that are most frequently mentioned throughout the articles. Out of these, overconfidence has garnered the most attention over the last decade, with 24 publications. Comparing the total number of articles with the number of articles per bias in Figure 4.2, it can be seen that many articles focused on more than one bias. Some of these articles presented reviews on the impact of biases, while others referred to relationships among certain biases. The following paragraphs will focus on the second, third, and fourth questions mentioned above and address the conceptualization of the biases, their roots in cognition and motivation, and their potential impact on

innovation-related decision making.

Representation and availability-related biases

Human agents can systematically misinterpret signal similarity if they fail to account for other relevant factors. They may relate an object's characteristics to an underlying category that does not truly represent the object [16]. For example, potential business partners might heavily overestimate the possibility of founding a successful start-up because their information environment may overstate success probabilities. The media may feature success stories more prominently and the founders could ignore base rates that imply significantly higher failure rates.

In our literature search, we identified seven articles that referred to misinterpretation of signal representativeness as the origin of bias. The studies applied laboratory experiments [20][25][43][74] or field studies with individual subject data [16][42] to explore susceptibility to representation-related biases.

It is not surprising that all studies referred to cognitive heuristics as an origin for the bias, as representation-related biases directly stem from the non-motivated misinterpretation of environmental information. One of the studies investigated whether genetic factors could explain susceptibility to management-related biases such as representativeness biases. In a twin-based field study, the researcher found evidence that cognition-related biases such as representativeness- and availability-related biases can be increased by genetic factors [16].

Three of the experimental studies and one field study specifically focused on forecasting problems, which we have closely related to innovation-evaluation in the previous chapters. In the first example, a series of experiments investigated how subjective probability estimates systematically depend on partitioning the variables that are being estimated [25]. A popular approach in assessing subjective probabilities is to transform (continuous) outcome variables into sets of exclusive and exhaustive events, e.g. "Will the new smartphone sell less than, equal to or more than x units?". The experimenters found that "assessed probabilities are systematically biased toward a uniform distribution over all events into which the relevant state space happens to be partitioned" [25]. Domain expertise may reduce but not eliminate such dependency of forecasts on outcome-variable partitions. The researchers' advice to provide forecasters with multiple partitions is to raise awareness that forecasters may assign probabilities based on partitions independent of the underlying prediction objects. The other two experimental studies focused on the

reaction to diagnostic signals such as past demand figures. The researchers found that forecasters underreact to signals in unstable environments and overreact to signals in stable environments because they ignore the underlying systems or base rates [43]. This relationship is attenuated if the stable environment shows less true underlying trends (i.e., “permanent shocks in the time series”[43]). Furthermore, forecasters, who are very good at predicting extreme outcomes, are sub-average predictors overall. Precisely predicting the potential of a radical yet highly successful innovation might actually be signal of poor judgment, because (a) extreme events such as highly successful radical innovations are very rare and (b) forecasters who take into account all available information would be less likely to predict such rare outcomes [20]. A field study in the domain of movie success predictions using information markets further indicates base-rate neglect. If forecasters have personal preferences for certain prediction objects, predictions of these objects will likely be overly optimistic. For example, subjects will overgeneralize their personal preference for sophisticated movies and underestimate the potential for non-elitist genres such as thrillers [42].

The last experimental study focused on the perception of innovative products [74]. Subjects were presented with either incrementally new or radically new products. The researchers found that potential customers will often evaluate radically new products as considerably worse than incrementally new products because they fail to imagine their use benefits. Yet when subjects are provided with several examples of use benefits, radically new products will receive better evaluations than incrementally new products due to the fact that contextual (overly represented) factors such as difficulty experienced during the visualization process are less important [74].

Overall, we find that individuals will often fail to validly access and assess diagnostic information to form evaluations. The forecasting experiments particularly highlight human difficulties in drawing valid conclusions from statistical information such as distribution of past demand or the performance of alternative products. The last experiment described shows that future use value may be particularly hard to evaluate because it remains a virtual and thus underrepresented entity in the present compared to diagnostic signals such as potential switching costs. The findings imply the importance of assisting individuals when they are searching for and using information to form beliefs in the context of innovation evaluation.

Confirmation biases

Confirmation biases build on the observation that people will more likely seek or interpret information in accordance with their existing beliefs, expectations or formulated hypotheses (Nickerson 1998). This bias has been closely related to positive test strategies that incline subjects to assume a statement to be true in the absence of compelling evidence for or against it. Individuals do not naturally attempt to falsify their own hypotheses to validate their truth, but rather search for supportive arguments for these hypotheses (Evans 1989).

Five articles from our literature sample addressed confirmatory information seeking and interpretation. Three of articles took a formal approach in addressing confirmation biases among entrepreneurs [6][10][62]. The first article explored why entrepreneurs and business founders attribute much of their entrepreneurial decision making to intuition. The authors particularly argued for limited cognitive ability as a bias origin. Company founders often apply positive hypothesis testing to attribute venture decisions to their *intuition* because they “[are] simply not able to consciously identify any more verifiable, obvious or compelling basis for having proceeded with the venture founding”[10]. The second article found that entrepreneurs appear especially prone to confirmation biases. Sole founders have more opportunity to attribute positive signals to personal decisions in the absence of status-competing peers, who may frequently provide contrary evidence. For example, in a corporate environment, managers’ assessments are more likely to be directly challenged by colleagues, which may prevent managers from looking only at confirmatory information. The researchers propose that entrepreneurs who are less susceptible to extant confirmation seeking are more likely to be successful [6]. Accordingly, confirmation seeking may cause entrepreneurs to delay business failure. They escalate their commitment in business ventures because alternative action would require them to falsify the initial hypothesis of venture success. In the article, the researchers argue that confirmatory information seeking for delaying business failure may allow founders to balance emotional distress, since longer periods of “anticipatory grieving” may lower the level of grief triggered by the failure event and allow them to recuperate faster [62]. Apart from entrepreneurship, confirmatory information seeking may also impact organizational learning ability. A qualitative study showed that confirmatory information seeking may prevent organizations from learning from rare negative events such as failed innovation projects. As these events provide singular points of evidence, their causes are often attributed to external factors before related data and implications are fully

exploited. This is especially true for organizations with successful histories in similar undertakings. Here, true learning from negative rare events is curbed by seeking information from past success, which supports the attribution of failure to external factors [67].

A final experimental study concluded that managers are perceived to be more effective and better at evaluating their own performance if they actively seek positive and negative feedback from superiors, peers, and subordinates [2]. Individuals perform better in organizational contexts if they overcome confirmation bias and show willingness to personally challenge positive self-image. Such a finding may be especially helpful in organizational evaluation tasks where individuals must consider heterogeneous information, such as in the evaluation of innovation.

In conclusion, confirmatory information seeking may strongly influence individual decision making by biasing them toward support of previously rendered beliefs. Three of the five studies reviewed particularly highlight that decision quality will be positively related to the ability to seek information that is relevant but that does not systematically support a previously formed hypothesis. In the context of innovation evaluation, confirmation biases may block openness to new information or evidence that contradicts previously-made personal assessments.

Loss Aversion

In total, 16 of the articles addressed loss aversion (LA). LA is a central component Prospect Theory (PT), which was developed by Kahneman and Tversky (1979). PT challenges the classical expected value theory for decision making under risk, in which expected values are rendered on the basis of a normative axiomatic foundation [14]. Instead, PT suggests that individuals perceive the value of risky decisions relative to reference points, so that marginal utility decreases with distance from the reference point [60]. Subjects overestimate the probability of unlikely events [41]. In this context, LA biases individuals to weigh losses and disadvantages more so than gains or advantages [41]. Losses loom larger than changes for the better [54]. As a consequence of LA, PT predicts that subjects will prefer riskier choices when alternative decision options are presented as sure losses, than if they are expected to yield sure gains [66].

The selected articles more frequently explained LA more as having motivational rather than cognitive origins (10 vs. 3). On the individual level of corporate decisions makers, the pronounced motivational rationale for loss aversion generalizes from behavioral

4. *Judgmental Biases in Innovation Evaluation*

ecology, which stresses “instincts for switching risk preference in the face of survival”: Human agents risk more not only to avoid death but to avoid any certain loss [14][16][31]. Cognitive awareness and experience with risky situations counteracts the motivation to avert losses [1][31]. On a firm level, higher risk taking in loss situations is explained by the motivation to keep the business afloat [30], especially if certain losses would result in total failure [62].

We found three main types of studies in the review: theoretical analysis [1][14][41][62], laboratory experiments [31][47][54][60][66] and field studies at the firm level [18][30][40][47]. On the firm level, the studies explored the impact of actual financial situations, as compared to aspiration levels, on risk-seeking behavior [40][47] and innovation-related spending [18][30]. All studies found support for LA. Firms invest more in innovation when they are performing below aspirational levels. This is especially true for family-owned businesses, where losses are more directly related to the personal wealth of decision makers [30]. Hence, perceiving organizational performance as below target may induce firms to reach overly favorable evaluations in order to start innovation endeavors.

The experimental studies aimed at exploring the relationship of LA to risk taking in more detail. In particular, the likelihood of making riskier investments in loss situations appears positively related to the degree to which subjects experience unpleasant feelings (through increased loss expectation) or pleasant feelings (through mood maintenance) [60]. Innovation managers are more likely to invest in highly innovative (and perceptually more risky) ventures than in incremental innovation projects in the face of losses from previous investments [58]. Additionally, framing subjects with a positive history in risky decision outcomes attenuates risk propensity, which positively impacts risky decision making in loss and gain situations [66]. Compared to described outcomes of risky decisions, self-experienced outcomes increase pessimism with regard to outcome probabilities, but only for gains and not for losses [31]. Subjects only learn to adopt LA in gain scenarios. Finally, LA affects real option valuations. Real option problems closely resemble information markets as mechanisms for uncertain corporate decision making. LA induces buyers and sellers to price options below their expected values because call options can be framed as alternatives to sure gains and put options as alternatives to potential losses [48]. Researchers stress that the aggregate evaluations of multiple risky decisions may yield more valid results because it may mitigate the effect of overly pronounced risk-aversion in gains and risk-seeking in losses [41].

To sum up, the studies demonstrate that LA contributes significantly to explaining situations that impact the likelihood of innovation undertakings. As an integral part of

prospect theory, LA especially affects instances in which innovations may be spurred by perceiving riskless alternatives as sure losses. This can be important in the context of innovation evaluation, because it may require evaluators to assess whether environmental sure loss conditions incline them to be overly risk-seeking and biased towards investing in innovation objects.

Status-Quo Bias

The status-quo bias describes a human tendency to disproportionately choose options that are consistent with the previous course of action (Samuelson and Zeckhauser 1988). Status-quo biases have been closely related to and even rooted in the previously described phenomena of endowment effects and loss aversion. Tversky and Kahneman (1991) found status-quo biases consistent with loss aversion, as proposed by Prospect Theory when past actions have led to success. In a situation where previous actions have resulted in success, a change of action will be unlikely to be preferred, as the reference point (continuing the course of action) would likely be perceived as a sure gain. Yet, status quo biases even persist when there are no gain or loss frames, in which case the bias is unlikely to be solely prompted by loss aversion (Samuelson and Zeckhauser 1988). Ritov and Brainerd (1992) show that status quo may be at least partially triggered in cases where maintaining the status quo requires no action and changing it requires action. Their experiments indicate that status quo biases do not materialize when keeping the status quo and changing it requires equal action and that status quo biases can be reversed in cases where only keeping the status quo requires action.

In total, 11 articles in the sample referred to status quo biases. Out of these, only a few articles referred to the psychological foundations of the bias. One article that studied genetic sources of various management-related biases found that status-quo-consistent decisions may be traced to genetic disposition [16]. Three articles specifically addressing the circumstances in which status quo biases occur in top management decision making argued that successful past performance acts as a cognitive anchor that prevents changing the course of action [4][11][66], even if new circumstances allow the decision maker to do so. Past engagement and success in risky decisions may increase risk propensity [66], decrease risk perception [66], and disengage the search for critical opinions, which is closely related to confirmatory information seeking [4]. Based on official records, one field study showed that the longer the tenure of top managers and the higher their past performance, the less likely they will be to commit to strategic change [11]. In another

field study that focused on official CEO statements, the researchers documented that CEOs who pay more attention to future events will be quicker to detect and develop innovation opportunities [73]. Yet, the more innovative the undertaking, the less likely the executives will be to abandon the innovation project in the light of poor forecasts, most likely because they expect excessive payouts if they complete it [58].

Compared to employed managers, entrepreneurs appear to be less susceptible to status quo biases because they are, by definition, individuals who have self-selected into challenging the status quo by starting businesses [12]. Yet, in a way closely related to overconfidence and confirmation biases, entrepreneurs may exhibit status quo biases in the context of failing ventures [21][62]. One experimental study among entrepreneurs showed that, similar to the case with top managers, previous venture success will increase status quo bias, as will personal investment or the lack of personal options besides the venture [21]. These personal drivers are, however, moderated by the entrepreneur's degree of extrinsic motivation. The more an entrepreneur seeks financial reward, the less these factors will hinder him from accepting business failure [21].

In summary, status quo biases are particularly relevant in contexts where innovations compete with on-going concerns regarding resources within organizations. Especially in organizations that have experienced long-term business success in the past, status quo biases may negatively impact the evaluation of innovations that systematically part from previous business practices, such as in the introduction of radically new technologies or very different customer groups.

Overconfidence

Researchers commonly describe overconfidence as “inaccurate, overly positive perceptions of one's abilities or knowledge” (Anderson et al. 2012). Yet, Moore and Healy (2008) highlight that the literature has often referred to three distinct facets of overconfidence that reflect the findings from our sample: **(1) Overestimation of actual performance (often coined *optimism*), (2) overplacement of relative performance, and (3) overprecision regarding outcome certainty.**

The first variety of overconfidence describes the overestimation of absolute performance, level of control, and chance of success, such as in overestimating the ability to successfully develop an innovation. According to Moore and Healy (2008), overestimation is the most empirically studied kind of overconfidence. The second type of overconfidence refers to believing oneself to be better than others at tasks such as evaluating the potential success of an innovation. The last kind of overconfidence expresses exten-

4. Judgmental Biases in Innovation Evaluation

sive certainty in the precision of one's beliefs. Such overprecision has repeatedly been demonstrated in experimental studies in which participants were typically required to provide 90% confidence intervals to questions with numerical answers, but where these intervals frequently comprised less than 50% true answers (Soll and Klayman 2004).

Moore and Small (2007) show that overplacement and overestimation are positively related if subjects draw inferences about relative placement by first learning about the performance of other subjects. This positive relationship particularly emerges in hard tasks that more closely resemble tasks in innovation management.

These findings may resemble the reality of confidence formation in innovation evaluations tasks. Camerer and Lovallo (1999) posit that overplacement results from perceived gaps between task difficulty for oneself and task difficulty for others. The authors argue that neglecting reference groups frequently yields lower expectations about personal capability and likelihood of others to perform comparably well on tasks. For example, innovators that aim to venture into new technology fields will often underestimate the likelihood of competition in such fields.

The entrepreneurship literature highlights that individuals will often form perceptions of relative performance by rating their own *potential* performance as superior to the performance of others (Mobius et al. 2011). Signals about others' past innovation success are more available than cues about others' failures and personal performance in innovation projects that will only generate results in the future. Diagnostic signals about one's own performance in entrepreneurial and innovation-related tasks are hard to come by because outcomes of entrepreneurial activity always lie in the future for newly-founded companies (Simon and Houghton 2003; Hayward et al. 2004). Furthermore, if this information is available but unsatisfactory, it may be neglected, supporting the notion of confirmatory information search (Hayward et al. 2006).

We found 33 studies in our sample that referred to overconfidence. Out of these, 26 articles identified motivation as the important driver for overconfidence, most notably self-enhancement and self-attribution motives. 16 articles referred to cognitive drivers, of which many closely related overconfidence to the misinterpretation of signals, (e.g. [35][41][65]).

The vast majority of our sample investigated overexpectation as a manifestation of overconfidence. 19 articles investigated the relationship of optimism to innovation-related tasks. Out of these, six articles specifically focused on firm actors such as innovation managers [65] or top managers [9][28][32][36][44] and concluded that overconfidence is

4. Judgmental Biases in Innovation Evaluation

positively related to the pursuit, but not the success, of innovation projects. Ten articles focused on entrepreneurial tendencies to be overly optimistic. Facing similar uncertain situations, entrepreneurs are more confident about the opportunities that stem from uncertainty and their ability to exploit these opportunities [6][34][41][45][55]. The studies confirm the findings of Cooper (1988), who stress that such success certainty is systematically unrealistic among business founders. A more recent simulation study found that self-selection effects may play an important role in the overoptimism of individuals in entrepreneurial roles [38].

Four articles positively related the tendency to overestimate personal capabilities to overplacement of performance in relation to others [32][41][55][70]. Overplacement has been found to negatively impact information seeking when starting a business, an effect that is attenuated in more uncertain domains [19]. Such a relationship may be explained by increased myopia in areas that are more unfamiliar [52].

Overprecision was the focus of four articles. Those studies found a general tendency toward overprecision in judgment [13][46][63], which may be more pronounced in more uncertain domains [13]. Overprecision can be successfully reduced by providing valid feedback or stimulating estimators to consider alternative answers [46]. One recent experiment demonstrated that advisors' confidence in advice precision may negatively affect advice seekers' payoffs [57].

In summary, overconfidence appears to be a very present bias in the context of innovation-related tasks in organizations. Overestimation of oneself and overplacement in relation to others seem to be pronounced in environments that feature high degrees of uncertainty. In such environments, outcomes are less likely to be traced back to personal performance, which supports attributive or neglectful behavior in positive or negative outcomes. The domain of entrepreneurship comprises the bulk of the literature. It may be easier to regard and study new firms or their founders as potentially overconfident entities because overconfidence and success variables relate more closely to each other in new firms. In contrast, established firms may attract overconfident individuals only in particular uncertainty- and innovation-related departments, which may complicate the study of the relationship between overconfidence and organizational success.

4.4. Judgmental biases in information markets for innovation evaluation

Information markets for innovation evaluation can be prone to biases that may have negative impact on their results.

In researching biases in information markets, it is possible to draw from specific observations of **financial asset markets** that are not necessarily related to evaluating innovations. We highlighted in the previous chapter that information markets are closely related to financial or asset markets, which underlines the presence of market-related biases in information markets. While specific research on the impact of biases in information markets is still scarce (Wu et al. 2008), we can draw on the large body of experimental and field research in financial markets.

One early study tested the extent to which markets can effectively aggregate information (Plott and Sunder 1982). Information was usually provided in an unambiguous and well-defined form, e.g. as precise likelihoods. The researchers could not refute the notion that privately distributed information is efficiently reflected by market prices in these environments. In reality, however, no market aggregates objective information, but rather, subjective beliefs. Markets are made by agents that act within the limits of bounded rationality (Camerer and Lovo 1999) and real market environments often provoke biased actions because information hardly allows for unambiguous processing and application.

In the context of innovation evaluation, information market participants may be subject to the very biases that affect individuals in innovation management, as were presented in the previous section. The underlying rationales for initiating and running information markets for innovation evaluation reflect the motivations that were discussed in the previous sections. While the information market resembles a novel mechanism for innovation evaluation, initiators and participants are comparable to any situation, in which innovations need to be evaluated. We can therefore integrate findings from the literature with observations from specific conditions in information markets and discuss the impact of biases that are subject to innovation evaluation tasks rather than market mechanisms.

Representation-related biases influence asset prices and prediction errors in information markets. In field experiments at Google and HP, potential prediction outcomes were partitioned into multiple stocks, where each stock represented a certain interval band

4. Judgmental Biases in Innovation Evaluation

of the outcome space (Chen and Plott 2002; Cowgill et al. 2008). Yet, another series of laboratory experiments showed that (1) partitions have significant effect on final market prices and that (2) partitions similarly affect experienced traders or market experts (Sonnemann 2008). Initiators' partition decisions may negatively impact the ability of participants to reveal valid information. "Unpacking one event (of three) into two component intervals increases its judged probability by about .25" (Sonnemann 2008). Closely related to the representativeness bias discussed in Section 4.3.2, subjects base their information market actions on the amount and concrete boundaries of events that underlie the stocks. Sonnemann (2008) suggests that initiators may reduce potential biases and improve prediction quality by running simultaneous markets that feature different partitions and aggregate results from these markets.

Furthermore, participants in (experimental) asset markets have been frequently found to confuse trading to maximize portfolio values with trading to increase the price of preferred stocks. Subjects do not cease from buying preferred stocks, even if they receive strong signals that these stocks will likely perform weakly. Research commonly refers to this phenomenon as the "wishful-thinking effect". Forsythe et al. (1999) used experimental markets to explore the wishful-thinking effect in political information markets. If subjects have external incentives that motivate high stock prices (e.g. a high vote share of the preferred political party), their trading will be biased towards market-external preferences. Seybert and Bloomfield (2009) ran a set of experiments to discriminate between wishful thinking (trading) based on observing market signals that manifest into larger bets on preferred outcomes, and overestimation that is based on overly positive private beliefs. Their results indicate that market-based wishful thinking will frequently be contagious, causing the private overestimations of other participants. As subjects usually learn about each others' beliefs via bets (or market trades), they "start to think wishfully." In the experiments, market trading appeared to positively accentuate the wishful beliefs of a few traders via the market-based wishful thinking of many.

Such an observation can be particularly important in the context of information markets for innovation evaluation. In this case, participants will likely relate the positive outcome of certain stocks with potentially larger personal benefit. For instance, potential customers may gain more utility from one potential product characteristic than another, and employees may benefit more if information markets predict positive outcomes for objects that were developed in their own departments. In these examples, stocks would include a price premium based on overestimation and wishful-thinking.

Ultimately, overconfidence must be considered in the context of information markets for innovation evaluation since certain potential participants are particularly likely to be overconfident and because it has often been observed that overconfident subjects can negatively influence the outcome of the related asset market.

Section 4.3.2 has shown that overconfidence appears more prominent among business founders, entrepreneurs, innovation managers, inventors and those top managers who are more strongly engaged in developing and introducing innovations. Additionally, the previous sections highlighted that financial markets may particularly attract overconfident market agents and suffer from their participation. The last decade has provided vivid examples of how a few apparently overconfident individuals can bring enormous losses to very large institutions (Clark 2008). As such, researchers have demonstrated genuine interest in evaluating the impact of overconfidence on asset markets (e.g. Malmendier et al. (2011)).

Two streams of research in the context of asset markets have investigated the potential impact of overconfident traders on individual gains and market price efficiency. One line of research extended classical market models with agents who sacrifice rationality for overconfidence regarding the precision of their information. In these models, overconfident investors overestimate the value of private estimations, hold risky portfolios, trade excessively, and deter stock market prices from reaching their fundamental value (Odean 1998; Daniel et al. 1998; Gervais 2001). Additionally, empirical results have revealed that overconfidence significantly impacts asset trading. Most prominently, overconfident subjects show increased trading activity and lowest individual performance in current findings (see Wu et al. (2008) for a review). Closely related to information markets, Deaves et al. (2009) studied the impact of different manifestations of overconfidence on trading behavior in an experimental asset market. The authors found that overestimation and better-than-average effects are positively correlated and significantly increment overall trading activity in an experimental asset market. Yet their experiments do not provide insight on how overconfidence impacts overall market efficiency. Wu et al. (2008) stress that more research relating individual overconfidence to market outcomes is needed.

4.5. Summary

This aims of this chapter were two-fold: first, to provide a more thorough understanding about the origins and impacts of important judgmental biases in the context of innovation and its evaluation; and second, to discuss the impact of judgmental biases on

information markets in the context of innovation evaluation.

With regard to the first target, we found that rich evidence exists demonstrating a detrimental impact of judgmental biases on the success of innovations. With regard to the origin of biases, the results lack clarity. While many articles related the biases in question to motivational or cognitive sources, we can not assert that researchers have attributed the biases to unequivocal origins by mutual consent. It appears that more research is needed to better discriminate (or unite, compare Kunda (1990)) and explain the origins of judgmental biases in the context of innovation evaluation. Based on the literature review, we found five biases that were most prominently featured. Their characteristics and impact in the context of innovation evaluation have been presented in detail above. Out of these, overconfidence garnered the most attention by researchers in innovation management. The review shows that overconfidence can be considered a driving force behind much innovation activity. Overconfidence is frequently found amongst entrepreneurs and managers who particularly embrace innovation. Overconfident subjects are more likely see a benefit in engaging in innovation and in reaching positive evaluations of innovative undertakings because they underestimate their risk of failure.

Focusing on the impact of judgmental biases on information in the context of innovation evaluation, the current research in the domain of financial markets has provided important insights for future research. We found that previous research from analogous domains has explored and analyzed the impact of various judgmental biases on mechanisms and situations that can be closely related to information markets for innovation evaluation. The findings show that important biases in the context of innovation evaluation, such as endowment effects, optimism or overconfidence, frequently and significantly impact agents' behavior and outcomes in markets that are closely related to information markets.

In sum, judgmental biases will often have strong influence on decision-making in the context of innovation evaluation. The literature review reveals that most studies have attributed negative consequences to the presence of judgmental biases. The review of the research has also suggested that information markets may suffer from participants' judgmental biases, even though these markets have been deemed highly efficient in aggregating information. However, very little empirical insight exists regarding how information markets function when exposed to judgmental biases in the context of innovation evaluation.

Part III.

Empirical Studies

5. Research Framework

In the following chapter, we will carve out the structural elements of our empirical research. On the one hand, the previous chapters have illustrated the potential benefits of applying information markets as organizational mechanism to source and aggregate information for innovation evaluation. On the other hand, we have discussed how innovation evaluation in general and information markets in particular may be prone to judgmental biases. This chapter will consider these findings in the context of our research goals. First, we will discuss a research focus on the overconfidence bias in the context of information markets in Section 5.1. In Section 5.2, an appropriate level of analysis is introduced. Next, we will continue by identifying a suitable research method in Section 5.3. Lastly, Section 5.4 will highlight the research process and introduce our empirical study.

5.1. Focus on overconfidence

This thesis set out to provide a more profound understanding on how judgmental biases impact innovation evaluation. Our have findings so far demonstrated that a considerable number of biases can be identified in the context of innovation management and information markets. Section 4.4 illustrated that information markets for innovation evaluation can frequently incorporate innovation-related biases; however, we also found that current direct insights are scarce and often anecdotal in nature. Systematic observation is required to substantiate an understanding of the potential effects of bias.

As Franke (2002) stresses in the tradition of Karl Popper’s critical rationalism, a researcher needs to test theory-based hypotheses via empirical investigation to reveal objective *realities*. Based on the findings gathered from the literature, we therefore need to develop assumptions about the effects of particular biases on the outcomes of innovation evaluation tasks via information markets and test these assumptions empirically. However, available resources, the vast array of biases, and their distinct effects require us to limit our empirical focus.

5. Research Framework

Based on the findings presented in Section 4.3 and Section 4.4, we chose overconfidence as the focal bias for our empirical research.

Overconfidence is a very relevant and prominently-featured bias in the context of innovation management and (financial) market research. The previous section demonstrated that overconfidence is prominently featured in the research of both domains. More specifically, overconfidence has been shown to have potentially negative influence on the outcome of financial markets' efficiency in aggregating information (Odean 1998) and on the success of innovation endeavors (Hayward et al. 2006). As information markets draw from market efficiency to increase the viability of innovation evaluation (and hence, the success potential of innovations), gaining a profound understanding about the impact of overconfidence in this context is imperative.

Overconfidence has been particularly observed among individuals who are frequently present in the context of innovation evaluation. Entrepreneurs (Trevelyan 2008), inventors (Astebro et al. 2007) and innovation-advocating executives (Galasso and Simcoe 2011), for example, have all been found to exhibit above average degrees of overconfidence and are likely participants in information markets used to evaluate the potential of innovations.

So far, very limited research exists that specifically focuses the impact of overconfidence on the efficiency of information markets as instruments for innovation evaluation. While Chapter 3 highlighted the fact that information markets have garnered increasing attention as an effective tool for evaluating innovations, Chapter 4 showed that limited research exists in the specific domain of innovation evaluation regarding how overconfidence affects results when applying information markets to evaluate the potential of innovations. While previous research has specifically called for empirical research to study the impact of overconfidence in information markets (Wu et al. 2008), these calls have so far gone unanswered.

Follow upon the recent work on overconfidence in psychological experiments that was presented in Chapter 4, we conceptualize overconfidence in the context of innovation evaluation as the belief of having an unrealistically good understanding about the future potential of an innovation (Anderson et al. 2012). Overconfident subjects draw expectations about their performance and others' performance from skewed distributions, which leads them to overestimate their own performance and overplace themselves among peers (Larrick et al. 2007). Accordingly, our concept of overconfidence will manifest itself in two ways: the overconfident subject will overestimate his ability to evaluate the potential

of innovations, and the overconfident subject will overplace his performance in relation to peers who also participate in the group-based evaluation task.

Theoretical work in the context of entrepreneurship has indicated that such a differentiation between perceived personal performance and the performance of others is common in innovation-evaluation tasks. Hayward et al. (2006) argue that although founders are aware that most ventures fail, they believe that they will beat the odds. We therefore focus on the impact of overconfidence that leads individuals to overestimate innovation-evaluation performance on its own and in relation to others.

5.2. Level of analysis

Overconfidence is an **individual characteristic**. However, based on the findings discussed in Chapter 4, the impact of individual overconfidence on innovation-evaluation at the group level can be driven by two interrelated domains: the impact of overconfidence on overconfident subjects' individual actions and the interaction of overconfident subjects with other participants.

First, the characteristics and influence of overconfidence can be observed on an individual level. Overconfidence is a trait that finds expression via overconfident subjects' actions. Overconfidence acts upon decision making particularly in the context of innovation evaluation. Confidence appears to be **variably pronounced** among individuals who participate in innovation evaluation tasks. While mild overconfidence in one's own personal ability in evaluation tasks has been found to be a very common human trait by early psychological studies (Fischhoff et al. 1977), considerably higher and more heterogeneous degrees of overconfidence appear to be particularly present in innovation evaluation tasks (Hayward et al. 2006). We presented arguments in the previous section that self-selection may systematically attract some individuals with particularly high degrees of overconfidence, such as entrepreneurs, inventors and top-managers, to participate in innovation tasks. The previous chapter showed that the degree of individual overconfidence influences individual evaluation of innovations, but we also stressed that an understanding of this phenomenon in the context of information markets is lacking. Hence, the first level of analysis in our study addresses the genesis of individual overconfidence in innovation evaluation and its impact on individual decision making in innovation evaluation via information markets.

Second, overconfident subjects interact in information markets with other participants. On the one hand, overconfidence could potentially influence the trading by overconfident

subjects, which would then be perceived, processed and incorporated by other participants into their actions. On the other hand, overconfidence could potentially influence the way overconfident subjects perceive other participants' behavior in the market. For example, as discussed in the previous chapter, overconfident subjects may be less open to market signals to update private information (Deaves et al. 2009). Both directions of interaction can be studied by uniting multiple subjects in information markets, which leads to the second level of analysis.

In summary, to build an understanding about the impact of overconfidence on the individual level, yet also gain valuable insight about its impact on the quality of information markets for innovation evaluation, we need to study the particular influence of overconfidence on (1) individual behavior and (2) interactions in information markets. Only then can we gain a true understanding about the relationship between overconfidence, individual action and information market outcomes.

Our empirical studies therefore need to focus on two distinct levels of analysis. On the first level, we will explore the impact of overconfidence on individual behavior in information markets **in a controlled individual environment**. On the second level, we will investigate the impact of overconfidence on the evaluation quality of an information market **in an integrated environment using individual agents but group-based outcomes**.

5.3. Methodology

We focus our study on the impact of the individual characteristic of overconfidence on both individual behavior and the group-level outcome (i.e. evaluation quality) of information markets. Sorensen et al. (2010) show that surveys, field studies, and case studies have traditionally been the methods of choice for innovation management researchers, although researchers have stressed that laboratory experiments would provide more valuable results if the research object “[were] practically and meaningfully isolated from a broader context of innovation” (Sorensen et al. 2010). In our research, overconfidence acts as the independent variable, while the dependent variables are individual behavior in information markets and innovation evaluation quality. Our research object therefore fits the abovementioned suggestion for applying laboratory experiments.

In psychology, causal relationships between traits and behavior are predominantly stud-

5. Research Framework

ied via laboratory experiments (Shaughnessey et al. 2012). It has been suggested that similar research questions in the field of management decision making should employ the same methods as psychology, although experiments could suffer from low contextual realism (Scandura and Williams 2000). Laboratory experiments can provide significant benefits in answering our research questions.

However, significant **practical barriers** oppose field studies. Our subjects cannot be easily sampled or recruited for field experiments. Furthermore, our focal characteristic, overconfidence, is heterogeneously dispersed and is not easily observed among individuals (Anderson and Kilduff 2009). Additionally, overconfidence appears concentrated among individuals like CEOs, venture capitalists and entrepreneurs, who presumably have relatively little time to participate in extensive studies that focus on their individual behavior and interactions in group tasks. Furthermore, overconfident yet high-ranking subjects may be more reluctant to participate in experiments that could ultimately undermine their decision-making competency.

More importantly, focusing on the **scientific methodology**, we can only draw inferences about causal relationships between confidence and individual actions if we successfully control for confidence levels among subjects. The strength of laboratory experiments lies in the rigid control over variables and environmental conditions in order to draw conclusions about such causal relationships (Willer and Walker 2007). The systematic manipulation of overconfidence in a laboratory experiment produces three advantages for our research approach.

First, laboratory experiments allow the recruitment of **homogeneous subjects** in order to reduce variance in characteristics other than the manipulated feature. However, it should be noted that a large degree of variety in subject characteristics provides additional variance and unwanted error when relating overconfidence to individual behavior and information-market outcomes.

Second, laboratory experiments allow researchers to **exert the best amount of control over the desired independent variable**. A successful manipulation creates variance in individual confidence that allows the valid investigation of the relationship between (over)confidence and behavior in information markets. Aside from its practical barriers, manipulating confidence outside the laboratory may likely be impossible because individuals are highly influenced by the context and cues given in previous tasks. Experiences of task performance within their natural environments are embedded within individuals, which likely reduces the potential effect of confidence manipulation.

Third, laboratory experiments can create a **close relationship between the source**

of overconfidence and the focal task, in which overconfidence influences subject behavior. Research has shown that individual levels of overconfidence do not necessarily span domains. Being overconfident in one task does not reliably predict overconfidence levels in another task. Perceived task difficulty and social context largely predict the relevant facets of overconfidence. Tasks that subjects perceive to be difficult for themselves but even harder for others likely produce the highest amount of overconfidence, both in an absolute sense and in relation to others (Larrick et al. 2007). To induce the least amount variance when studying overconfidence in innovation evaluation tasks, it may therefore be important to manipulate overconfidence via a close relationship between the source of overconfidence and the task domain.

Overall, a critical assumption underlying the interpretation of laboratory experiments is that the **insights gained in the lab can be generalized to the world beyond**. For physical laws and processes, evidence supports the idea that what happens in the laboratory is equally valid in the broader world. Yet much controversy exists regarding the conditions in which the same will hold true for social experiments that focus on individual psychological conditions or the behavior of human groups. Levitt and List (2007) highlight that moral and ethical considerations, the nature and extent of scrutiny of ones actions by others, (self-)selection processes, the context of decision making, and the stakes of decision making heavily influence behavior in the lab, apart from monetary considerations. The researchers stress that these factors differ between the laboratory and the social context the lab aims to resemble, which could prevent researcher from being able to generalize findings in a straightforward manner.

Other researchers have discussed recruiting **student subjects** in particular. Recruiting students brings the advantage that the recruiting process does not require extant resources, as subjects are readily available and usually cheap. Another advantage is that student characteristics can be better controlled in terms of educational background and domain knowledge, compared to drawing subjects from the general population. Thus, results would be more closely related to the experimental manipulation since they would exhibit less unexplained error. Additionally, carefully selected subjects may be more appropriate than random samples in cases where the experimental results are to be related to professional groups that can hardly be recruited for experiments. For example, accounting MBA students have been found to be appropriate subjects for studying the behavior of professionals, while bachelor students in economics did not sufficiently resemble accountants' behaviors in various experimental tasks (Liyanarachchi 2007). Similar findings have also been reported for studies related to the underlying research of group-

5. Research Framework

based innovation evaluation, e.g. the finding that undergraduate students should not be used to study the behavior of managers in decision support systems but that participants can be appropriately sampled from graduate students (Remus 1989). As a consequence, recruiting students may uncover systematic error if researchers falsely assume the generalizability of the student behavior in the experiment to the real-world behavior that the researcher aims to explain.

In the end, Levitt and List (2007) emphasize the exclusive value of laboratory experiments that study economic behavior, regardless of the abovementioned shortcomings. Instead of using experimental results as highly diagnostic cues for real-life behavior, the authors suggest to understand them as a crucial first understanding regarding fundamental behavioral mechanisms that would otherwise be difficult to observe and understand in real-world studies. Our experiments build upon this suggestion.

Our empirical studies address the impact of the individual characteristic of overconfidence on individual behavior in information markets and on innovation evaluation quality as a group-based outcome of information market trading. **Separate experiments** allow us to isolate these two levels of analysis and their respective goals to reach a more valid conclusion regarding the impact of overconfidence on each dependent variable. As our analysis units are individual behavior in information markets and outcomes of information markets, we selected individual and market-related outcomes as granular dependent variables.

On the individual level, we needed to develop a controlled environment that allowed us to isolate and investigate the relationship between overconfidence and individual behavior. On the level of information market outcomes, we needed to test information market performance in the light of real innovation evaluation tasks to validly assess the impact of overconfidence on information market efficacy.

Finally, it would be difficult to recruit a sufficient number of subjects when focusing on outcomes of single market periods as the granular level of analysis. Each market period characterizes only case for both levels of analysis.

Many experiments in behavioral market economics have encountered such a subject-related bottleneck by allowing subjects to participate in more than one market period while still analyzing each period independently (e.g. Seybert and Bloomfield (2009), Healy et al. (2010), and Jian and Sami (2011)). **Within-subject designs** are generally accepted in psychological and economic experimental research, but researchers have

been encouraged to better account for subject-based errors during analysis. Methodological work has stressed the importance of applying multi-level approaches when analysis requires controlling for within-subject error (Judd et al. 2001) or aims to explain group-level outcomes (Kashy and Kenny 2000; Van Kleef et al. 2010).

To summarize, our empirical studies use laboratory experiments to investigate the impact of overconfidence in innovation evaluation by drawing from student subjects. One experimental study is focused on individual behavior in information markets, while the other experiment is focused on the impact of overconfidence on the outcome of information markets for innovation evaluation. Subjects interact in multiple market periods, a condition which has been successfully controlled for by using within-subject designs to analyze the results.

5.4. Research process

We constrained our research focus to the impact of individual overconfidence in information markets and, as illustrated in the previous section, the experimental methods chosen appear to be most salient for studying the individual and group-based level of analysis. Figure 5.1 reflects our research process for the empirical studies. We began at the most granular level by testing whether confidence in innovation evaluation could be experimentally manipulated as an individual trait. After developing a feasible manipulation to create artificial confidence levels in the laboratory, we continued on to the first experiment addressing the relationship between confidence and participation behavior in the context of information markets. Here, the focus was set upon individual behavior in a controlled market environment to isolate the effects of confidence. Finally, we studied the interaction between individual participation behavior at controlled confidence levels in information markets and other human agents. In this experiment, the relationship between participants' confidence levels and information market evaluation quality became the focus of attention.

In the following subsections, we will introduce the three steps of the research process. Each subsection will briefly address the theoretical and empirical background for our study and highlight our specific goals with regard to the gaps in the current literature and our central research questions.

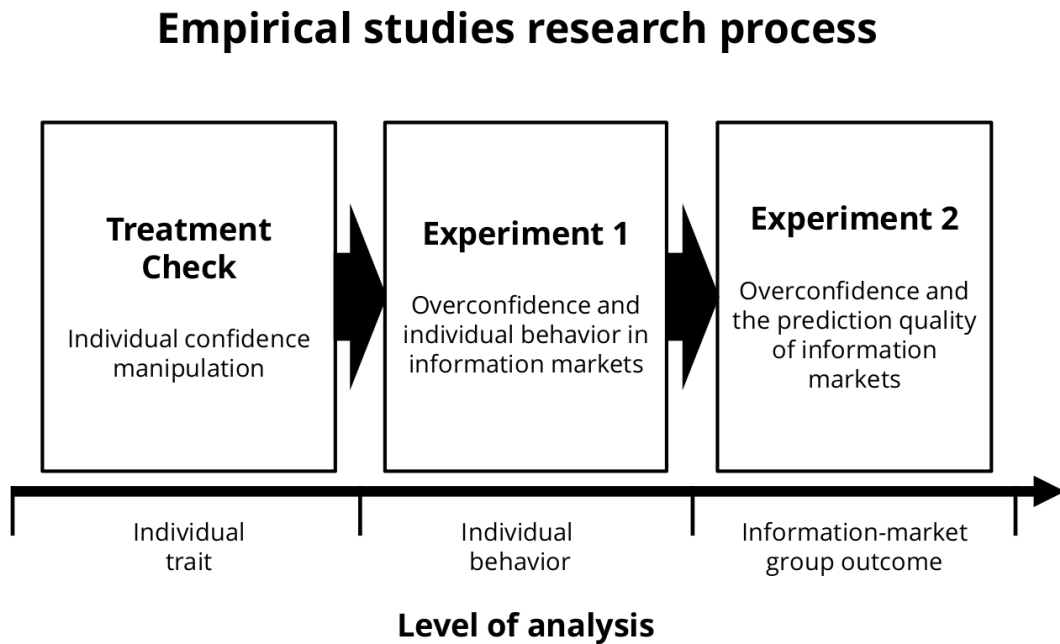


Figure 5.1.: Research process and level of analysis of the three empirical studies (Source: Own depiction)

5.4.1. Individual confidence manipulation

Causal relationships are studied in laboratory experiments by examining the effects that result from manipulating a particular impact factor. In our case, we aimed to study the impact of different confidence levels.

Drawing on research in social psychology, there are numerous examples where experimenters have successfully shaped participants' self-concepts via **false-feedback manipulation**.

Frey (1978) provides an early example in which personal performance is experimentally manipulated to test subjects' reactions to success and failure in public or private conditions after receiving or evaluating performance feedback. In the experiment, student subjects took an intelligence test and received fictitious results that were either below or above the average. Here, the performance-feedback manipulation induced significant differences in subjects' subsequent behavior. Depending on their perceived performance and public or private feedback on their test scores, the subjects systematically chose different strategies to increase presented self-value.

More closely related to self-concept manipulation, Greenberg et al. (1992) artificially

induced high and low self-esteem via false feedback. These experiments focused on the relationship between self esteem and anxiety levels in experiencing and expecting shock. The authors found that student subjects manipulated to have high self-esteem physiologically experienced and reported significantly lower levels of anxiety when experiencing or expecting shock.

The specific **manipulation of overconfidence** has only recently been applied in a psychological experiment. Anderson et al. (2012) studied the status-enhancing effect of overconfidence. The researchers used overly positive feedback on a pre-experimental estimation task to induce overconfidence. This treatment successfully induced overconfidence in student subjects while preventing subjects from gaining increased self-esteem or suspecting that they had been given false feedback. After the manipulation, overconfident subjects were paired with subjects who did not receive the treatment to repeat a similar task as a team. The study showed that artificially-induced overconfidence positively influences subjects' perceived competency and status, both by themselves and by their respective partners. However, the researchers stress that the confidence manipulation carried out specifically focused on the ability to predict other peoples' characteristics. They highlight that individuals are often largely unaware of their accuracy in perceiving others, which makes it easier to exploit in experimental manipulation (Ames and Kammrath 2004). While their study gives indication that confidence can be successfully manipulated in a laboratory, it lacks relation to overconfidence in the context of innovation evaluation.

In summary, a large body of experimental research in psychology has demonstrated the feasibility and success of manipulating self-concepts in general and overconfidence in particular. Based on the research, it is possible to develop and test a suitable treatment that allows us to induce overconfidence by providing manipulated feedback. The treatment check addresses the domain of innovation evaluation because we later focus on innovation evaluation tasks. Also, the above-mentioned study by Anderson et al. (2012) highlighted that the specific manipulation they used should not be indiscriminately transferred to different task domains.

5.4.2. Overconfidence and individual behavior in information markets

On the individual level, we aimed to gain an understanding of the impact of overconfidence on individual behavior in information markets. However, in real-life information

5. Research Framework

markets with multiple human participants, observing individual traits and overall market behavior to help understand relationships between individual traits and individual behavior may be impossible. All market actions (except the first market action) are potentially subject to interactions between the initial subjective condition and the market actions of other participants. This would likely prevent a researcher from validly explain variance in a subject's behavior based on his condition alone, which, in our case, relates to confidence. Yet, researchers have also stressed that too little is known about the direct relationship between confidence and market behavior (Oberlechner and Osler 2012). Our first experiment therefore set out to create an experimental environment that would allow us to learn about the direct effects of different confidence levels on individual trading behavior in information markets.

In an experimental setting, a researcher should rigorously control the environment across different treatment conditions. In a market experiment, such control requires stability of the market environment across treatments. Researchers have previously engaged actors to control and stabilize live human interactions in studies that focused on economic behavior (e.g. Kopelman et al. (2006)). In an information market, such actors could aim to keep their trading stable based on pre-defined rules that are remain constant over experimental conditions. Yet, information market experiments do not necessarily require the visual presence of other human participants. The human actor may be exchangeable with an artificial agent if it acts sufficiently human. Earlier research has pointed to the importance of human subjects at least perceiving the agents to be human. Otherwise, they could exhibit behavior that might be more geared toward interaction with artificial machines, e.g. appearing to be less disciplined in their actions (Brown-Kruse 1991). Yet, more recent experiments in market economics have provided evidence that artificial agents can be incorporated in human-subject experiments without provoking unnatural human-subject behavior (Duffy 2006). We drew from these findings to create an artificial market environment for the first experiment that would best control the impact of overconfidence on individual behavior. In the experiment, human subjects would perceive that they were trading with other human subjects, while really interacting with an artificial market maker.

We integrated the findings from Section 4.3 and Section 4.4 to develop our hypotheses about the relationship between overconfidence and individual behavior in information markets for innovation evaluation. The hypotheses will be presented at the beginning of Chapter 7.

5.4.3. Overconfidence and the prediction quality of information markets

The experimental design of the first experiment prevented us from drawing inferences about the impact of overconfidence on the prediction quality of information markets. However, prediction error is arguably the most important success variable in the context of information markets for innovation evaluation. Initiators of information markets will be more interested to learn how to deal with excessive confidence if it significantly alters the predictive quality of the markets. Hence, we developed a second experiment, presented in Chapter 8, that allowed us to study the impact of overconfidence on prediction quality as a group-based outcome of information markets. We integrated the theoretical findings from Section 4.4 about the potential impact of overconfidence on the prediction quality of information markets for innovation evaluation and also drew from the empirical findings that will be reported in Chapter 7. Here, particular attention will be given to the dimensions of increased aggressiveness in trading by overconfident subjects and the contingency effects of this aggressiveness on non-treated subjects' trading and evaluation behavior.

Furthermore, Section 2.2 documented that innovation projects likely induce different behaviors for searching out and acquiring diagnostic information, depending on how much information is initially available to the evaluators and the price to search for additional information.

We took these earlier findings into account by studying the impact of overconfidence on the prediction quality of information markets in two scenarios in which information market participants have distinct access to diagnostic information.

First, we created a basic experimental setting in which we controlled how well subjects were informed regarding the prediction task. All subjects received free pieces of similar diagnostic information, which was aimed at creating similar knowledge about the prediction task among the subjects. The subjects were expected to use the given information to form expectations regarding the prediction task and to trade within the information market. In this basic setting, we began by treating a set of subjects either with overconfidence or low confidence. Next, we awarded them with diagnostic information, and brought them together with a matching number of uninformed traders, who received neither manipulated feedback nor diagnostic information. Lastly, we compared the impact of overconfident traders compared to low-confidence traders on the prediction quality of the information markets in the case where treated subjects were given free access to

diagnostic information.

In addition, we investigated how overconfidence impacts the willingness of subjects to engage in costly information searching. For example, entrepreneurship scholars have shown that overconfident individuals are less likely to engage in information acquisition or search to make more reasoned decisions (Trevelyan 2008). As a consequence, higher confidence may negatively impact information seeking in innovation endeavors where uncertainty is particularly high. Again, a group of subjects were either treated with overconfidence or low confidence. Then, instead of receiving information for free, subjects were given the opportunity to acquire information for a cost after learning about the information markets' underlying prediction task. As in the basic experimental setting, treated subjects then entered the information market together with a matching number of uninformed traders, who received neither manipulated feedback nor diagnostic information. Lastly, we assessed the impact of overconfident traders compared to low-confidence traders on the prediction quality of the information market in the case where the treated subjects were given the opportunity to acquire diagnostic information at a cost.

6. Treatment Check

We needed to develop and test a feasible treatment for manipulating individual confidence before we could study the impact of different confidence levels under laboratory conditions. In order to introduce **overconfidence as an experimental variable**, randomly selected subjects needed to exhibit overconfidence overall and relative to their peers after treatment. This chapter will present the treatment check used. We tested the feasibility of our confidence treatment in an experiment conducted prior to the information market experiments. In this chapter, the experimental design and the implementation of the treatment check will first be presented in detail. We will then discuss our choice of innovation evaluation tasks before showing that overconfidence can indeed be induced experimentally.

6.1. Experimental design and implementation

We aimed to induce overconfidence and low confidence by presenting the subjects with **manipulated feedback** following ten innovation evaluation questions. We conducted a treatment check with 99 graduate and undergraduate engineering students at Hamburg University of Technology, who were recruited in an introductory business course lecture. Subjects were randomly assigned to one of three experimental groups: overconfidence, low-confidence, or the control group, which did not receive any treatment.

The treatment check was carried out by each subject on individual computer workstations. It consisted of **two rounds of ten evaluation questions and one feedback interval** for the three experimental groups, as depicted in Figure 6.1.¹ The treatment was issued after the first round of evaluation questions and is henceforth described as either positive or negative feedback.

After a brief explanatory screen, the experiment began with the first round of evaluation questions. All questions were related to innovative developments in the German auto-

¹The evaluation questions can be found in the appendix.

6. Treatment Check

mobile industry. In each question, the subjects needed to estimate percentage values between 0 and 100, e.g. the percentage of newly registered hybrid vehicles in Germany in the current month. The subjects were told that estimates would be regarded correct if they differed from the true underlying value by no more than ten percent. We choose this interval so that subjects would truly believe they had performed exceptionally well in the case of positive feedback.

The initial ten-question task was completed after the subjects had estimated how many questions they believed they had answered correctly. For each round of evaluation questions, they were offered the chance to win a €50,- voucher for a large online retail store. They were told that the chance of winning would be equally influenced by how many questions they answered correctly and by how well they estimated their own performance. Subjects were told that it was in their best interest **to correctly answer as many questions as possible and to evaluate the own performance as well as possible**.

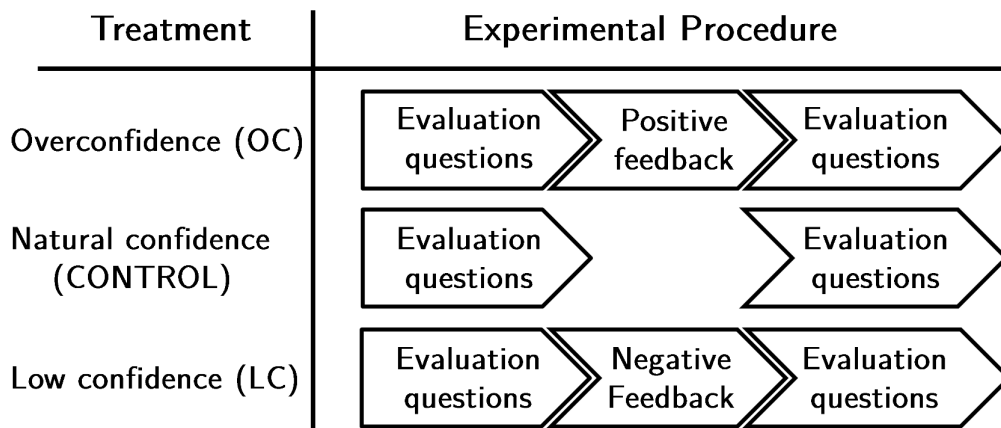


Figure 6.1.: Experimental procedure, by treatment group, in the treatment check
(Source: Own depiction)

When confidence is well calibrated, subjects should be able to guess the number of questions answered correctly. In the case that subjects believed they had answered more questions correctly than truly answered correctly, they were deemed as naturally overconfident, and if subjects answered more questions correctly than they had indicated, they were categorized as naturally underconfident.

After the first round of questions and the self-assessment, **subjects in the treatment groups were presented with manipulated feedback, indicating that they answered eight questions correctly if they were in the overconfidence group, or eight questions incorrectly if they were in the low-confidence group.** Subjects were provided with a manipulated histogram that displayed their alleged relative performance among their peers in the experiment. Subjects in the control group did not receive any feedback. Afterwards, all subjects were asked to answer an additional set of ten questions, including a self-assessment to measure the manipulation effect.

6.2. Innovation evaluation tasks

We chose innovation-related prediction tasks from the **German automobile industry** for a number of reasons.

First, the automobile industry is highly relevant and by far the most attractive future employer for German engineering students (Trendence Institut 2011). Thus, we hypothesized that the interest and involvement of our subjects would be sufficiently high to make informed predictions regarding specific real-world evaluation tasks. In addition, master students in the engineering disciplines are more likely to be found among managers in engineering companies than students from other disciplines.

Second, German public offices for automobile registration statistics are required to frequently publish information that allowed us to validate the true underlying values of our tasks and hence incentivize our participants according to their true information market performance. We developed a set of 20 innovation-related questions ranging from assessing the potential market success of specific models to evaluating the market share development of drive train technologies, e.g. the market share growth of hybrid cars among newly registered vehicles.

Our evaluation tasks differed from typical tasks in behavioral experiments since we incorporated **real-world innovation evaluations** in our experiment. We decided to abandon the economic experimental paradigm and to use abstract stimuli in our information markets for two important reasons.

First, we built upon the previously-discussed theory that overconfidence stems largely from the misinterpretation of task relevant information. We believed that **memorized information** is accessed very early and perceived to be highly diagnostic for rendering evaluations. Memorized information thus heavily influences the willingness to use

6. Treatment Check

Variable	Low confidence	CONTROL	Overconfidence
Confidence level in the first evaluation task	1.97 (2.13)	1.78 (2.36)	1.79 (2.17)
Confidence level in the second evaluation task	-0.06 1.47	0.94 2.54	2.13 2.01
Change in confidence			
between first and second evaluation task	-2.03 (1.82)	-0.84 (2.11)	0.34 (2.25)

Notes. $n = 99$; MANOVA for between-group differences in confidence change is significant at $p < 0.001$; post-hoc LSD tests between groups within dependent variables are significant at $p < 0.05$.

Table 6.1.: Treatment check results

and adopt new information. We therefore required prediction targets that allowed participants to access their memory to generate initial beliefs. According to the theory presented, this will foster subjects' ability to attribute their performance to their superior evaluation skills.

Second, innovation evaluation is a relevant practical task in a specific environment. We aimed to shed new light on potential barriers to valid innovation evaluation. We were convinced that we could best balance **external and internal validity** by using real-world stimuli for our evaluation tasks. On the one hand, the arguments presented above highlight the importance of using stimuli that resonate with subjects' memorized information. On the other hand, a smaller bridge between real-world tasks and experimental tasks increases the validity of theoretical inferences about the impact of overconfidence in the specific domain of innovation evaluation and does not limit the results to the domain of experimental economics.

6.3. Results

The statistical results of the treatment check are presented in Table 6.1 and are further illustrated by Figure 6.2. We conducted a multivariate analysis of variance (MANOVA) to examine whether (a) the treatment led to **significant group differences in confidence** after the second evaluation task and (b) confidence was significantly higher in the overconfidence group than among non-treated and low-confidence treated individuals.

The table reveals that the treatment significantly differentiated the confidence levels between the three experimental groups ($F = 9.576$; $p < 0.001$). Overconfident individuals

6. Treatment Check

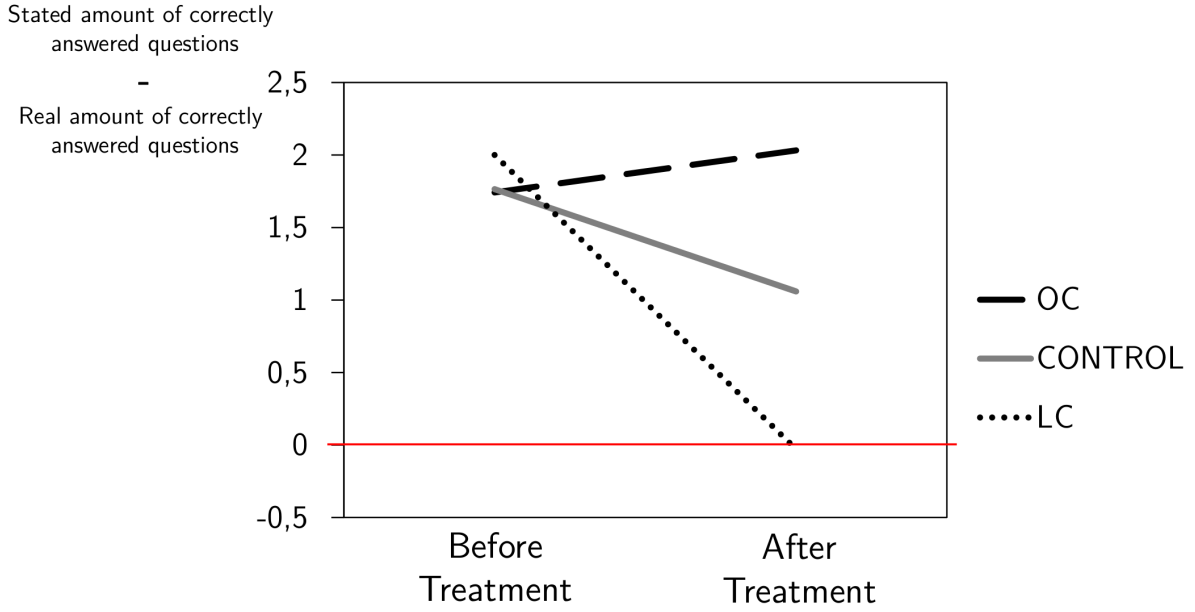


Figure 6.2.: Change in (over-)confidence before and after treatment in the treatment check (Source: Own depiction)

overestimated the amount of questions answered correctly ($\bar{x} = 2.13$) significantly more so ($p < 0.05$ for both pairwise comparisons) than members of the control group ($\bar{x} = 0.94$) and subjects in the low-confidence treatment ($\bar{x} = -0.06$). Additionally although the low-confidence subjects yielded the lowest confidence ($p < 0.05$), the treatment barely managed to make them under-estimate their evaluation precision. Figure ?? illustrates the change in confidence between pre- and post-treatment evaluation. All three experimental groups share roughly the same confidence before the treatment. We observe that overall confidence appeared to have sunk for the second evaluation after the treatment. This is additionally underlined by the change in confidence in the control-group. Questions may have been perceived more difficult to answer, which they were visibly not in reality. Figure 6.2 and the second row of Table 6.1 show that **changes in confidence following treatment differed significantly among the experimental groups** ($F = 11.064$; $p < 0.001$). **Only the subjects in the overconfidence treatment increased their confidence from pre- to post-evaluation** ($\bar{x} = 0.34$). Subjects in the control group ($\bar{x} = -0.84$) and the low confidence group ($\bar{x} = 2.03$) decreased in confidence between tasks, differing significantly from the overconfident subjects ($p < 0.05$ for both pairwise comparisons). Furthermore, subjects in the low confidence condition produced a significantly bigger drop in confidence than subjects in the control group

6. *Treatment Check*

($p < 0.05$). These results show that the treatment was successful in inducing higher and lower confidence in individuals with regard to innovation-related evaluation tasks. Overconfidence was significantly increased through our treatment and the negative feedback condition allowed us to significantly reduce confidence compared to the no treatment condition.

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

This chapter presents our first experiment, in which we studied the impact of overconfidence on individual behavior in information markets. By drawing from theoretical and empirical findings on the impact of overconfidence on individual behavior in the context of financial markets and innovation evaluation, three hypotheses were developed regarding how overconfidence would influence participants' trading in information markets. A laboratory experiment provided a controlled environment to aptly test the hypotheses and gain novel understanding on how overconfidence acts on individual information-market trading.

7.1. Development of hypotheses

Theoretical insight from psychology and empirical findings from closely related domains such as entrepreneurship, innovation management and financial markets suggest how overconfidence will impact individual behavior in the context of innovation evaluation via information markets. We can distinguish three behavioral effects that appear to follow overconfidence among investors, entrepreneurs, inventors and management decision-makers when making decisions in uncertain contexts such as financial investments, innovation evaluation or new business endeavors.

First, overconfident individuals seem to **make quicker decisions** than their non-overconfident peers (Statman et al. 2006). Second, they **pursue decisions with more aggressiveness** and vigor (Camerer and Lovallo 1999; Koellinger et al. 2007). Finally, overconfident individuals appear to be **less likely to update their beliefs** and modify their decisions when contradictory cues urge them to do so (Biyalogorsky et al. 2006; Astebro et al. 2007). In an information market, evaluations, beliefs, and the process of

individual belief formation and articulation can be observed through trading activity (Arrow et al. 2008). Hence, we ultimately rely on market-based variables to propose relationships between confidence and behavior in information markets.

Modeling decision making under uncertainty using **biased-hypothesis testing** integrates motivational and cognitive drivers and explains how decisions can be influenced by nuances in confidence (Feldman and Lynch 1988). Subjects approach decision-making under uncertainty with an initial hypothesis that is then tested in the light of additional information (Pyszczynski and Greenberg 1987). Such information is sought out sequentially based on its accessibility and is then evaluated for its perceived diagnostic utility. **Cognitive load** for processing additional existing private information will be lowest because it is most easily accessible and most closely related to subjective mental models. Hence, private information will be acquired and processed first for testing the hypothesis (Gentner and Stevens 2014). At the same time, diagnosticity of new information depends on its relatedness to previously accessed information, which acts as an anchor for the validity and diagnosticity of all subsequent inputs (Tversky and Kahneman 1974). The more closely the new information concurs with the previously accessed information, the more likely that it will be regarded as helpful.

As a consequence, initial information inhibits evaluators' ability to consider conflicting assessments when being confronted with new and contradictory information. Alternative outcomes are likely to be ruled out. Subjects will only draw from conflicting information to test their hypothesis if it appears unambiguous and truly diagnostic. According to Pyszczynski and Greenberg (1987), the more confidence subjects possess, the more *biased* their hypothesis testing will be. Confidence increases relative value of initial private information compared to external signals.

Motivational drivers may further drive decision-related confidence and explain the degree of bias when modeling decision making as hypothesis testing (Kunda 1990). The stronger the subjects' motives for making accurate decisions, the more thoroughly they will generate and test hypotheses in the light of unbiased information retrieval, processing and evaluation. However, as Section 4.2.2 has shown, subjects may have motives that are separate from accurate decision making. For example, the more the evaluators will personally benefit from an innovation, the more motivated they will be to attribute success potential to that particular innovation, compared to alternative new endeavors. Such motivation may arise from self-enhancement in the case that the innovation has been introduced by the evaluators themselves or that it increases job security or facil-

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

itates wage raises (Shepperd et al. 2008). Furthermore, positive self-image from strong confidence may motivate overconfidence in subjects. Self-deceit in believing oneself to be relatively more confident and capable than in reality reduces cognitive load in the portrayal of this image to the external environment (Von Hippel and Trivers 2011). Much experimental evidence exists to support the notion that motivational biases influence how subjects generate hypotheses, search for information and then evaluate those hypotheses (Larrick et al. 2007). For example, self-threatening hypotheses produce aversive states, which lead subjects to select less threatening hypotheses (Pyszczynski and Greenberg 1987). Subjects will seek information that prevents aversive states of arousal and supports self-enhancement. When given either success or failure feedback on a supposedly valid performance-assessment test, subjects will particularly engage in information search to understand their performance in cases where which all subjects received low scores but not in cases where only they received low scores (Pyszczynski et al. 1985).

A frequent observation is that **overconfident subjects appear to require less time to evaluate situations and make decisions** in the context of innovation evaluation. Researchers have argued that spotting and acting more quickly upon perceived opportunities is at the very core of becoming an entrepreneur or inventor (Lowe and Ziedonis 2006). Wally and Baum (1994) found that executives who scored high on tolerance for risk, use of intuition and propensity to act, took speedier decisions in strategic contexts such as deciding on new business endeavors.

However, a stronger degree of biased hypothesis testing may explain why overconfident subjects require less time to act upon their expectations. Overconfidence leads subjects to attribute higher validity to signals that support initial assessments. Subjects require less initial information and less additional supportive information to reach a threshold at which their degree of information allows them to make decisions and act upon them. The more confident subjects are, the more affirmative private and public signals will be that appear to support the initial hypothesis or the initial assessment. The environmental conditions of entrepreneurship and innovation development positively influence the selection of overconfident subjects who engage in biased hypothesis testing because limited information on technical feasibility and market conditions favors subjects who are able to make decisions with little information (Busenitz and Barney 1997).

As discussed, decision making in the context of information markets for innovation evaluation can be observed via trading (Arrow et al. 2008). The more quickly individuals decide that their beliefs about future states (in our case, an innovation's success

potential) differ from current market predictions, the earlier they will assume trading opportunities to exploit that gap. They require less information from external market signals to validate their expectations. Accordingly, they begin trading more quickly in the information market. Based on all of the factors discussed above, we formed the following hypothesis:

Hypothesis 1 *Individual confidence will be positively related to earlier trading actions.*

Particularly in the domain of financial markets, overconfidence has often been related to the intensity and aggressiveness of belief expression via the **amount of investment**. Individuals who assume that they possess superior information, knowledge or capabilities have been observed to enter markets excessively (Camerer and Lovo 1999), even if their expected payouts are negatively related to increases in market participation. By analyzing the trading behavior of over 35,000 households using a large brokerage firm, Barber and Odean (2001) found that the group of more confident investors traded 45% more than the less confident group, but that they suffered from a larger penalty on returns by this increased trading activity. While household investors exhibit higher degrees of overconfidence in investment decisions, experts such as pension fund managers still suffer from significantly overestimating their capability in evaluating the future returns of investments (Gort et al. 2008). These findings are reflected in related domains and among other experts, too. For example, companies with overconfident CEOs invest more and use more debt than less confident peers (Ben-David et al. 2007).

As discussed for the previous hypothesis, overconfident subjects require less time and information to validate private assessments even when new information arises. More confidence reduces cognitive expenditure for processing additional information to validate an initial hypothesis that a certain investment opportunity appears attractive. Poorly-defined external signals such as market-price changes appear particularly less salient than private information; thus, the perceived marginal utility of waiting or actively searching for external information and processing it will be lower for overconfident individuals (Zacharakis and Shepherd 2001). Fewer resources are required to evaluate investments and more become available for investing.

In addition, overestimating the validity of private signals reduces awareness of underlying uncertainty regarding the evaluation objects (Odean 1998). As a consequence, overconfident subjects are more certain that predicted outcomes will materialize in the future. This reduces the necessity to invest time and effort in mechanisms to diversify risk Barber and Odean (2001). For instance, uncertain future outcomes usually require investors

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

to balance portfolios and hedge against risk. Overconfident subjects feel less of a need to engage in costly optimization and allocation strategies. Instead, they are more likely to invest funds in outcomes they feel are more likely. For example, overconfident CEOs invest systematically larger proportions of private funds in their companies (Malmendier et al. 2011) and overconfident financial investors invest larger amounts of funds and hold less diversified investment portfolios (Wu et al. 2008).

Based on the previous discussion, confidence may be positively related to the overall level of investment in information markets. Investments are made via purchasing or selling information market stock. Higher confidence leaves individuals with more time and mental capacity to engage in trading. This leads to our second hypothesis:

Hypothesis 2 *Individual confidence will be positively related to total stock turnover.*

Ultimately, and highly important from the perspective of obtaining valid innovation evaluations, overconfidence may prevent subjects from altering initial evaluations and updating private information based on non-affirmative public signals. We found many indicators in related research domains that overconfident individuals will be less likely to change their beliefs, even if new information strongly challenges their current assessments. Financial market researchers have similarly argued that higher individual overconfidence will emphasize the value of existing private information compared to novel public signals that become available (Daniel et al. 1997). Financial investors' beliefs in initial estimates are fostered by market information that is positively correlated with these beliefs but unlikely to be disturbed by alternative market signals (Baron 2000). In the same vein, inventors have been found to continue developing their innovations, even when experts tell them to stop "wasting" more funds (Astebro et al. 2007). Entrepreneurs will continue unsuccessful development for longer times than established companies because they are more likely to be overconfident regarding their capabilities for succeeding with their ventures (Lowe and Ziedonis 2006).

The previous paragraphs haven't given an indication of why overconfident individuals will be less likely to change their beliefs by relating the decision-making outcomes of overconfidence to biased hypotheses. In the evaluation process, information signals are considered in the light of previously acquired and processed information. The more these signals concur with previously evaluated information and the more they support the initial hypothesis, the more will they be regarded for the purpose of updating private assessments. This relationship is increased with higher degrees of confidence. As a consequence, overconfident subjects will focus on private assessments when making

evaluations in the innovation context because external evaluations are more likely to provide critical feedback. Many studies of have found this myopic self-focus to exist in market-entry decisions (Moore et al. 2007; Simon and Shrader 2012).

Furthermore, neglect of external information appears particularly strong in the later stages of the evaluation processes. Having invested in actions following assessments enforces these assessments' perceived validity, which has been attributed to endowment effects that overemphasize the value of private assets or self-enhancements motives (Hammond and Keeney 1998).

In an information market context, higher confidence will therefore be related to a lower degree of learning from market signals and less trading in accordance with these market signals. Less belief updating occurs because higher confidence will manifest through more trading based on initial beliefs. Additionally, reluctance to update beliefs based on market signals will become apparent when comparing pre- to post-market beliefs. Higher confidence will reduce the impact of market signals on belief updating. We therefore hypothesize that reduced belief updating will be expressed via the following effects:

Hypothesis 3 *Individual confidence will be positively related to higher net turnover towards initial private estimates.*

Hypothesis 4 *Individual confidence will be negatively related to the willingness to change before-market estimates based on information market signals.*

7.2. Experimental design

7.2.1. Participants

We recruited **114 graduate engineering students** from a German university of technology to participate in the experiment. All participants held bachelors' degrees in engineering disciplines and were currently enrolled in a business administration course at Hamburg University of Technology. The average age was 24 ($sd = 2,47$) and 22% were females.

The experimental subjects were drawn from different classes than the pre-test, thus ensuring that no pre-test participants were present in the main experiment, minimizing the risk of knowledge exchange between pre-test and experimental subjects.

Similar to the treatment check, we recruited student subjects with backgrounds at the

intersection of business and engineering, as these subjects would be more likely to bear resemblance to participants in real-world innovation evaluation tasks (Moenaert and Souder 1996).

We issued **pre- and post-experimental questionnaires** to control for deviations in the subjects' predisposition regarding relevant criteria for participating in information markets for innovation evaluation. By drawing from existing scales, we checked the subjects' interest in participating in financial markets, their likelihood of engaging in risky financial investments, and their involvement with the product domain for innovation evaluation used in the experiment.

After the market exercise, we presented the subjects with an information market quiz developed by us to test their understanding of how to act in information markets and benefit from superior private information through trading.

Ultimately, **15 subjects needed to be removed** from the analysis. Twelve subjects were excluded, either because their quiz answers showed that they did not understand the information market, or because the trading data indicated that they were displaying erratic behavior (e.g. always setting pre- and post-estimation values to "100" but not trading accordingly). Three additional subjects iteratively clicked "buy" and "sell" more than 400 times during the market, which is possible in our market maker environment. They committed 30 times more trades than the average participant. Thus, they were excluded from the analysis as market outliers. After removing these 15 subjects, 99 subjects participated successfully in the experiment.

7.2.2. Implementation

The experiment was implemented in z-Tree, a software framework for programming economic experiments (Fischbacher 2007). The laboratory setup for the experiment is depicted in Figure 7.1.

Eight subjects participated in each session. Subjects drew a number that determined their workstation during the session. Each workstation was separated by non-transparent walls that prevented any visual interaction during the experiment with any other subject or the experimenter. In addition, subjects wore noise-canceling headphones during the experiment to minimize the impact of hearing other subjects trade on individual participation behavior. To further avoid cross-subject influences, we reduced PC interaction to the proprietary track-pads and keypads. We did so after becoming aware in the pre-tests that using (and therefore clicking) computer mice for trading created a strong audible sensation among fellow subjects. Sessions were either run as treatment sessions

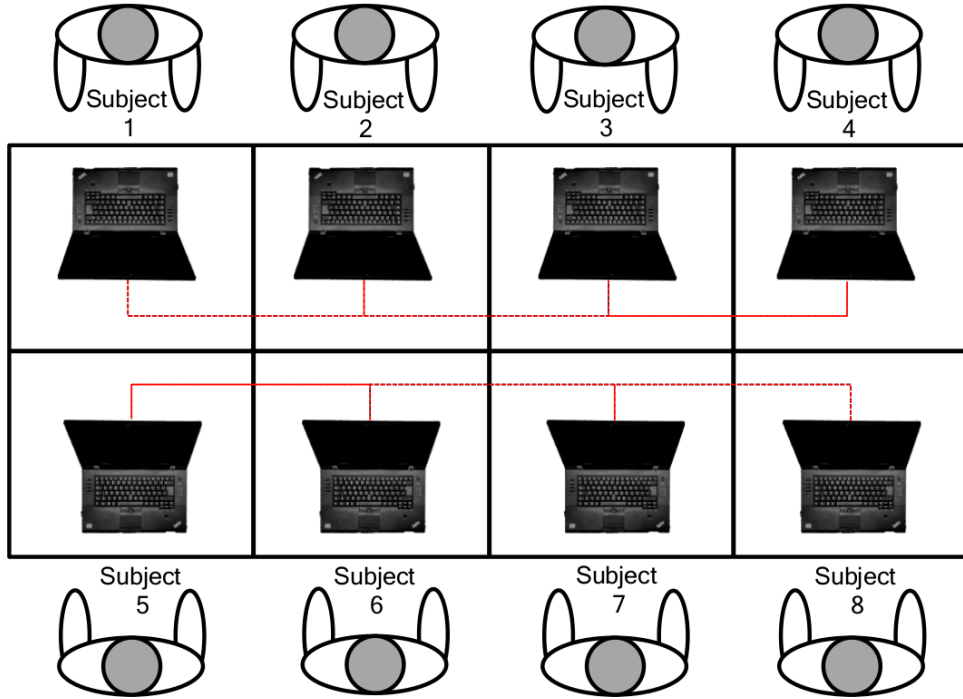


Figure 7.1.: Laboratory set-up for the first experiment (Source: Own depiction)

or control sessions. In treatment sessions, half of the subjects received an overconfidence treatment and half of the subjects received a low-confidence treatment.

While **subjects interacted in an isolated and artificial information market** where they did not trade with other human subjects, the experiment required them to believe they were interacting with other traders. Therefore, the control group subjects participated in separate sessions because their overall participation time was slightly lower than for treated subjects, as they did not participate in the feedback process. If they had participated jointly in sessions with treated subjects, they would have faced considerable waiting times during the feedback round to keep up the illusion of joint trading. The experiment consisted of five parts, as depicted in Figure 7.2.. At the beginning of the experiment, the subjects answered a pre-experimental questionnaire in which we checked for their involvement with the focal product domain and financial markets, and for their willingness to take financial risks, as potential control variables for later analysis. The subjects were then presented with the same ten evaluation questions given to the groups

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

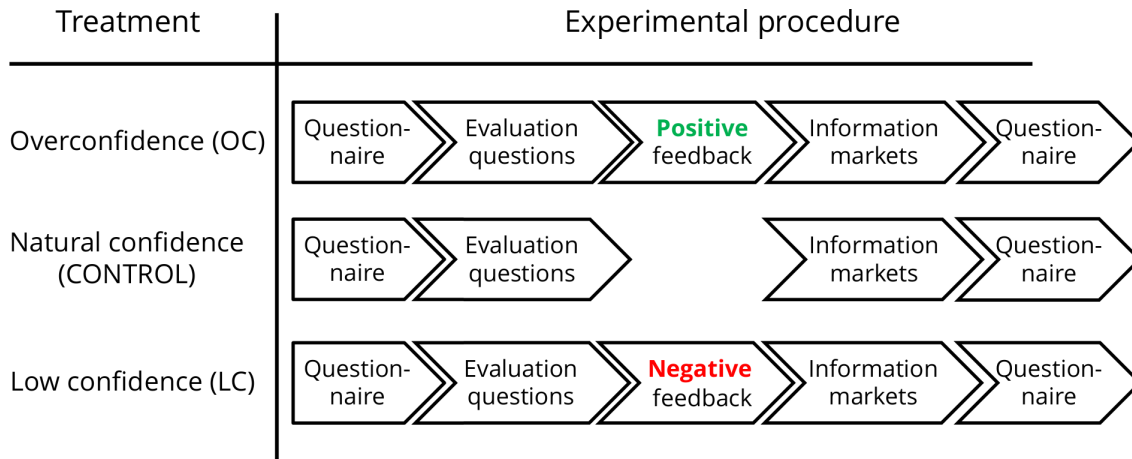


Figure 7.2.: Experimental procedure, by treatment group, in the first experiment (Source: Own depiction)

during the treatment check.¹ For the same reasons as stated for the treatment check, the innovation evaluation questions were drawn up based on novel developments in the German automobile industry. Again, subjects were told that their chance to win a €50,- voucher was influenced by how many questions they answered correctly and by how well they estimated their own performance. As in the treatment check, the manipulation was the third part of the experiment and remained unchanged, except that the histogram now compared the subject's performance with its alleged peers in the information market. Again, subjects in the control group did not receive any feedback.

The information market task consisted of **six independent market periods** following the feedback manipulation. Subjects were told that they would be trading with three participants in each market. Every market dealt with one innovation evaluation question, e.g. "By what percentage will the sales volume of electric vehicles increase in the current quarter, compared to the same period of the previous year?" Before and after trading, the subjects were asked to report their estimations of the true values to evaluate the effect of information market trading on individual belief updating regarding the true value. An additional monetary incentive was offered based on the mean error from these two estimates to incentivize the revelation of true beliefs.

Market running-time was considerably shorter than in most real-world applications of information markets used for innovation evaluation. Each market ran for 180 seconds.

¹The complete set of evaluation questions and information market prediction tasks are included in the Appendix.

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

This limited subjects' ability to gather novel information and learn new information external to the market. However, such a limitation was crucial to the experiment because our hypotheses were particularly aimed at studying overconfident traders' actions based on changes in the market-internal information environment, i.e. the market prices. Additionally, each market represented one case for subsequent empirical analysis. Stretching market running times to a duration that resembled real-world applications would have prevented a sufficient number of equally controlled cases that allowed for reliable statistical analysis. While the short running time may have therefore negatively influenced external validity, it was necessary to ensure internal validity and reliability.

We facilitated trading through an **automated market maker using Market Scoring Rules** to update prices (see Section 3.3.3 or Hanson (2003) for a detailed discussion of the mechanism). The experimental procedure incorporated means of incrementing the subjects' impression of a real inter-human information market. We derived the artificial agent's behavior by analyzing trading volumes and behavior in four-subject test markets before the experiment in order to simulate the action of three human co-participants.

Accordingly, the markets' starting prices were based on the subjective pre-market estimates but were not identical to the subjects' pre-market estimates. Instead they were 1) randomly set above or below the subjects' initial estimates and 2) randomly set at a distance of 10 or 25 percentage points from the subjects' initial estimates. For example, if the subject initially estimated a true underlying value of 30 percent, the market's starting price might have randomly been set at 55 percent, which is 25 percentage points above the initial estimate. Both means were used to create the impression of heterogeneous individual estimates of other market participants before the market.

If the market's starting price was set below a subject's initial estimate, the artificial agent traded towards a market price below the subject's initial estimate during the trading period. If the market starting price was set above the subject's initial estimate, the artificial agent traded towards a market price above the subject's initial estimate during the trading period.

The artificial agent was set to trade towards a market price that differed from the subject's belief by 25 percentage points if the starting price differed by 10 percentage points from the initial estimate, and by 50 percentage points if the starting price differed by 25 percentage points from the initial estimate.

The artificial agent's trading direction (towards a goal above or below the subject's initial estimate) and the goal of its trading (25 or 50 percentage points from the subject's initial estimate) were covered by appropriate covariates during empirical analysis.

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

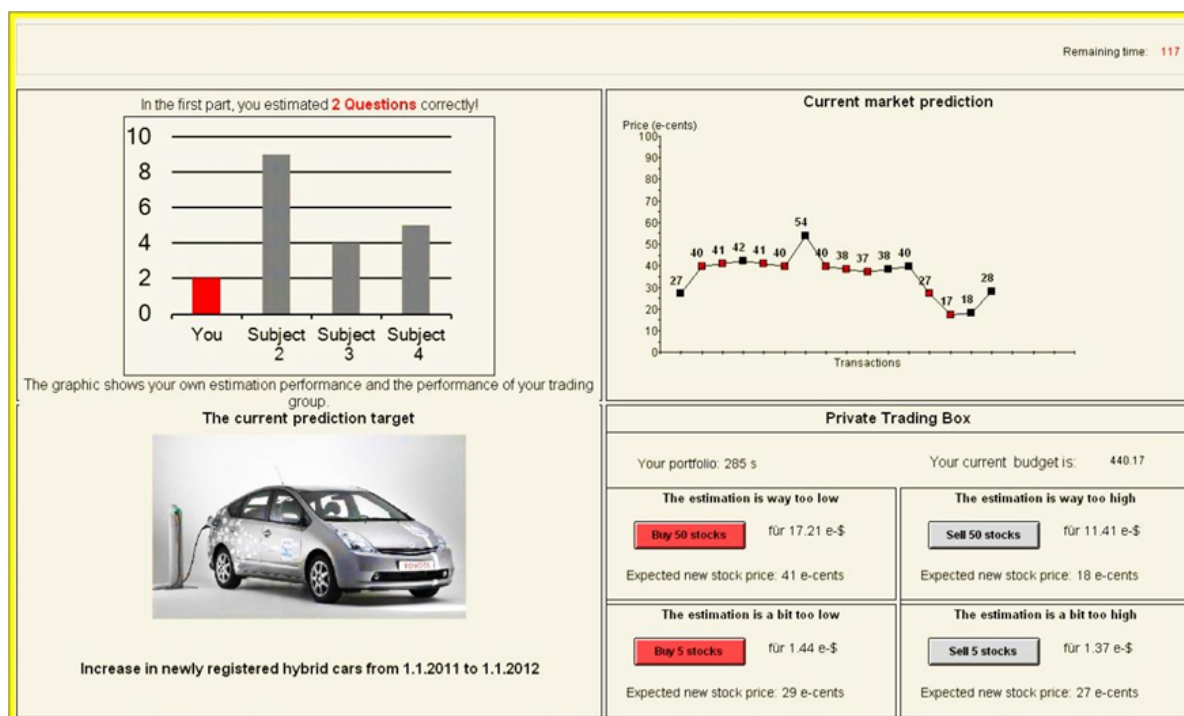


Figure 7.3.: Information market interaction screen for subjects under LC-treatment (Source: Own depiction)

Finally, a combinatorial design was chosen to arrange prediction tasks and periods over sessions and subjects. The design was adapted to cover eight sessions per treatment condition and to control for the effect of task and period in subsequent statistical analysis. Figure 7.3 shows the information market interface. The top left corner shows the treatment manipulation and the lower left corner the current prediction task. The stock price development is shown in the top right corner and participants can buy or sell either big (50) or small (5) packages of stock in the lower right corner, which also displays the current budget and portfolio.

After all markets had closed, two prediction tasks were selected randomly and the subjects' portfolios were paid out based on their true underlying value in the corresponding real-life market periods. Subjects were aware that the results would be determined by their performance in these randomly selected market periods. This procedure has previously been found to reduce the danger of carry-over effects, i.e. where perceived performance in one market period affects behavior in subsequent market periods (Deaves et al. 2009). Subjects were ranked based on the sum of portfolio values from the selected market periods. Their final rank determined their probability of winning two vouchers worth €50,- and €25,-.

The last part of the experiment consisted of the post-experimental questionnaire. First, we provided questions to evaluate subjective performance perceptions and subjective reasons for their respective performances. The questionnaire then asked subjects about how much they enjoyed the experimental tasks. Finally, the questionnaire addressed whether the subjects understood the information market correctly, and collected information on participants' age, gender, place of origin, and course of study.

7.2.3. Instructions

The day prior to their session, each subject received a written document with detailed instructions on all aspects of the innovation evaluation tasks (except the treatments), including the incentives, and how to trade and compete in the information market. Instead of including a practice session, we developed a thorough 10-minute video tutorial that was hard-wired into the experimental software. The video was displayed before the initial estimation task and guided the participants through all potential interactions of the evaluation task and information market. Two specific versions of the video were produced to fit the treatment and control groups. The control group version did not cover the feedback regarding the initial evaluation task.²

7.3. Sample and data preparation

7.3.1. Missing data

The experiment yielded data from 99 participants. 521 market periods were validly performed by the participants and recorded by the experimental software. It is important to explain the mismatch between the expected yield from the 99 participants, which would have been $99 \times 6 = 594$, and the actual yield, which was 521 market periods. Table 7.1 shows the number of subjects per market period yield. Out of 99 subjects, 63 subjects contributed all six market periods to the analysis. However, 36 subjects contributed only five or less periods to the analysis. There are two distinct reasons for this distribution. First, the subjects that only contributed one to three periods were part of the first experimental session, which was in fact the first real experimental session in the laboratory. While running flawlessly during pre-tests, the experiment server was evidently unable to handle the amount of subjects during the first session, which had eight participants. Only one subject (the one seated at the computer that was acting

²The version of the video for treated subjects can be viewed here: <http://youtu.be/yFl-GhfJ9yU>

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

Number of recorded market periods	Number of subjects
1	3
2	3
3	1
4	14
5	15
6	63

Table 7.1.: Number of subjects per recorded market periods yield

as the server) was recorded correctly, while periods from the other participants were not properly recorded. This became visible after the session. As a solution, another computer was set up as experimental server so that two separate servers could handle four subjects each. This setup recorded all sessions flawlessly thereafter. Second, the instances in which only four or five periods were recorded from subjects can be explained by programming errors. As discussed earlier, the starting prices in the markets were partly determined by the subjects' initial estimations. After the initial estimate, the market would either start below or above this estimate, either with a small or a big distance. In the erroneous sessions, however, the starting price was not taken from the corresponding initial estimate but from an estimate from a previous prediction target. Therefore, trade direction and trading distance were no longer related to the prediction task, as intended by the experimenter. This happened with two periods (periods two and five) in two sessions, after which it was corrected. It then occurred once more in two sessions with one period (period three). In total, the computer and programming errors caused $594 - 521 = 73$ or $\frac{73}{594} = 12.3\%$ **missing values** that needed to be accounted for and appropriately dealt with.

Two potential problems arise from missing data. First, missing data decreases the statistical power of subsequent analysis, which refers to the ability to discover significant relationships among variables in datasets (Roth 1994). The issue mainly becomes problematic if the given data shows no statistically significant relationships supporting the previously stated hypotheses and if the researcher holds a reasonable belief that the sample is too small. From a validity perspective, the second problem is far more concerning. Parameter estimates can be biased if data is systematically missing from specific parameters in variables that hypothetically relate to the dependent variables in question. If, for example, low-confidence treated individuals stopped evaluating their relative performance after the market because they were ashamed that their performance became

visible after the experiment, data would be systematically missing for a specific parameter of an important variable. Now, any form of substitution for the missing values would decrease correlation of confidence treatment and subjective performance evaluation, thus negatively impacting the results' validity (Roth 1994). It was therefore important to assess whether the data is Missing Not At Random (MNAR) (Allison 2001).

Participants' responses were not systematically skewed in our case, but the measurement was randomly interrupted and thus independent from all subjective responses and characteristics in the dataset. Yet, as the missing data was either generated in the first session or systematically during periods two and five, the data would need to be considered MNAR if the period or session numbers were significant indicators for other variables that are relevant to our analysis. Correlation analyses, however, which included both period number and session number, indicated no relationship of the two variables with any of the other variables in the model. We could therefore refute the MNAR assumption and consider the data Missing At Random (MAR), which allows us to ignore the missing data in subsequent analysis (Allison 2001).

7.3.2. Variable operationalization and construct validity

Before the data could be analyzed for its model fit and the hypotheses tested, they needed to be prepared and coded so that statistical results could be valid and be interpreted.

Table 7.2 presents the dependent and independent variables that are included in the statistical analyses and their code schemes. The upper part of the table shows the **dependent variables** and briefly describes their operationalization.

First, the overall stock turnover describes the total sum of information market stocks that were traded by subjects per market-period. The figure is calculated by simply adding up all of stocks that a participant traded in the period.

Second, the time of first trade describes the time difference, in seconds, between the start of a market period and the first trade by the subject. This variable is calculated by recording the time stamp at which the first trading action was executed by the participant. If a subject started trading immediately when the market started, the time of first trade for the period would be "0".

If a participant made a trade that resulted in a market price with a higher difference from their initial estimate, we added the stocks traded as "against initial estimate". *Vice versa*, if their trade reduced the difference between market price and their initial esti-

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

Variable	Operationalization by subject
Dependent variables	
Total stock turnover	Total sum of stocks traded
Time of the first trade	Seconds between the market period start and the subject's first trading activity
Net stock turnover towards initial estimate	The sum of stocks traded against the initial estimate subtracted from the sum of stocks traded towards the initial estimate by a subject in a market period
Delta between pre- and post-market estimate	The absolute difference between pre-market estimate and post-market estimate by a subject in a market period
Independent variable	
Treatment Condition	-1 (Low confidence) 0 (Natural confidence) 1 (Overconfidence)
Covariates	
Trading Distance	0 (Close distance) 1 (Far distance)
Artificial-Agent Trade Direction	-1 (Downwards) 1 (Upwards)
Gender	-1 (Female) 1 (Male)
Financial Product Risk Attitude	Metric coefficient (<i>extracted via regression from factor analysis</i>)
Product Domain Involvement	Metric coefficient (<i>extracted via regression from factor analysis</i>)

Table 7.2.: Operationalization of variables included in the statistical analysis for Experiment 1

mate, we added the stocks traded as “towards initial estimate”. By subtracting the sum of stocks traded “against initial estimate” from those traded “towards initial estimate”, we derived the net stock turnover towards the initial estimate. The delta between pre-market and post-market estimate is measured by the absolute difference between the individual prediction target estimations before and after each information market.

The **independent variable and covariates** were coded binary and did not require any preparation besides a sensible coding scheme.

The first covariate Trading Distance (TD) refers to the distance from the initial estimate that the artificial trading agent sets as its trading goal.

Artificial-Agent Trade Direction (AATD) describes whether the artificial agent has set his

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

trading goal higher or lower than the subject. When the artificial agent traded towards an estimate above the subject's estimation, the variable is set to "1" and "0" otherwise. Additionally, covariates for gender, Financial-Product Risk-Attitude (FPRA), and Product Domain Involvement (PDI) are included in the analysis. Gender is considered to be very closely related to the extent to which individuals are overconfident. It has previously been shown that men are more prone to higher confidence levels. A few researchers have gone so far as to use gender as a proxy for overconfidence (Barber and Odean 2001). The two construct-based covariates FPRA and PDI are included because they aim to control for domain familiarity in the experimental environment, which has two main domains: engagement and interest in risky financial trading; and exposure to new developments in the car industry. It has been frequently shown that domain familiarity can significantly impact subject overconfidence and subsequent behavior (Larrick et al. 2007). Among other aspects, subjects feel more certain in familiar domains and tend need less time to adapt to the environment, which supports the notion of controlling for FPRA and PDI. Both are established reflective constructs that are directly adopted from their authors (Weber et al. 2002; Chandrashekar 2004).

Items	Product Domain Involvement	Financial Products Risk Attitude
PDI1	0,879	0,142
PDI2	0,909	0,032
PDI3	0,621	0,300
FPRA1	0,156	0,636
FPRA2	0,081	0,839
FPRA3	0,121	0,632
Variance explained	0,34	0,27
Cronbach's α	0,76	0,59

Notes. Exploratory Factor Analysis: Principal Component Analysis (Eigenvalues > 1); Varimax-Rotation.

Table 7.3.: Measurement of constructs for covariates

Finally, the two construct-based covariates needed to be further investigated via factor analysis. First, it needed to be verified that the questionnaire items reliably reflected the underlying constructs, which was, in this case, indicated by sufficient levels of Cronbach's α for low item numbers (Peterson 1994). Carrying out a factor analysis using principal component analysis and varimax rotation, the rotated components matrix showed that all relevant factor loadings are higher than 0.6. As can be also seen in Table 7.3, cross loadings are very low and do not pass a threshold of 0.3, which would have called for fur-

ther investigation of discriminant validity (Ferguson and Cox 1993). The PDI construct explains 34% of total item variance, whereas the FPRA construct explains 27% of total item variance. Finally, the factor values for each subject were extracted via multiple linear regression to introduce individual construct values as covariates in the statistical analyses that followed.

7.3.3. Data adequacy and statistical model

The experiment used a **within-subject design** with multiple measurements per subject. Each subject participated in multiple market periods and each market period produced a single case. Dependent cases were nested within groups that shared independent variable characteristics, e.g. the subject's treatment condition or demographics. We therefore needed to consider the grouping of subjects' characteristics to explain variance between different subjects' market periods. Otherwise, sampling variance would be underestimated, which would lead to inflated alpha levels or type I error rates (Hox 1998). Significance tests ignoring the nested data structure would produce undue significant effects. Accordingly, the statistical analysis needed to be carried out using hierarchical linear models (HLMs). These can validly accommodate hierarchical data in which one response variable is measured at the lowest level (Level 1) while explanatory variable is measured at existing levels (Level 1+n) (Hox 1998).

In this experiment, market-period variables were entered into the models as their smallest units at Level 1 and subject characteristics were entered into the model as Level 2 grouping variables (Snijders and Bosker 2012).

We begin this section by displaying the descriptive results for the model variables. Ta-

Variable	Mean	sd	1	2	3	4	5
1. Total stock turnover	406.08	201.722					
2. Time of first trade	14.53	13.421	-0.223***				
3. Net stock turnover	275.70	231.775	0.489***	-0.303***			
towards initial estimate							
4. Delta between pre- and post-estimate	13.38	11.709	-0.140***	0.083 [†]	-0.028		
5. Trading Distance	0.48	0.500	0.109*	-0.041	0.276***	0.419***	
6. Artificial Agent Trade Direction	0.19	0.983	-0.123**	0.108*	-0.154***	-0.004	0.038

Notes. $n = 521$; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$.

Table 7.4.: Correlation, means and standard deviations for variables at Level 1

ble 7.4 shows the descriptive statistics for the Level 1 variables at information-market-

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

Variable	Mean	sd	1	2	3
1. Treatment	0.303	0.826			
2. Gender	0.434	0.905	0.197 [†]		
3. Factor FPRA	0 ¹	1.000 ¹	-0.064	0.298*	
4. Factor PDI	0 ¹	1.000 ¹	0.075	0.201*	0.002
<i>Notes.</i> $n = 99$; ¹ = Standardized regression coefficients; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$.					

Table 7.5.: Correlation, means and standard deviations for variables at Level 2

level based on the total sample of 521 information-market periods. The descriptive statistics for the subject-level variables are presented in Table 7.5. The table shows significant correlations between gender and both factor regression coefficients at the $p < 0.05$ level. However, checking variance inflation factors did not reveal any values larger than conservative cutoff points ($VIF > 10$, Belsley et al. (2005)) that may indeed indicate multicollinearity problems. The next step was to check whether the data was adequate for HLM analysis and fit model assumptions. We relied on the diagnoses suggested by prominent HLM-textbook authors (Raudenbush and Byrk 2002, pp. 252-286; Snijders and Bosker 2012, pp. 153-173). According to them, data adequacy should be checked and discussed by investigating fit using the following assumptions:

- Specified model distribution of outcome variables
- Normal distribution of Level 1 residuals
- Homogeneous variances of Level 1 residuals
- Normal distribution of Level 2 random coefficients
- Normal distribution of Level 2 residuals

The subsequent discussion of data adequacy is graphically supported by Figure 7.4, which displays the relevant histograms and scatter plots. We applied linear HLM models, which usually prescribe normal distributions for the four dependent variables (Raudenbush and Byrk 2002, p. 266). The first row in Figure 7.4 compares the normality assumption with the true distribution by comparing histograms of the sample distributions with unbiased normal distribution curves. We find that the two trading-amount-related variables fit nicely with normal assumption distributions, while the other two variables exhibit visible kurtosis (time of first trade = 4.271; Delta estimate = 1.079) and skew to the left (time of first trade = 1.815; Delta estimate = 1.067). While researchers have

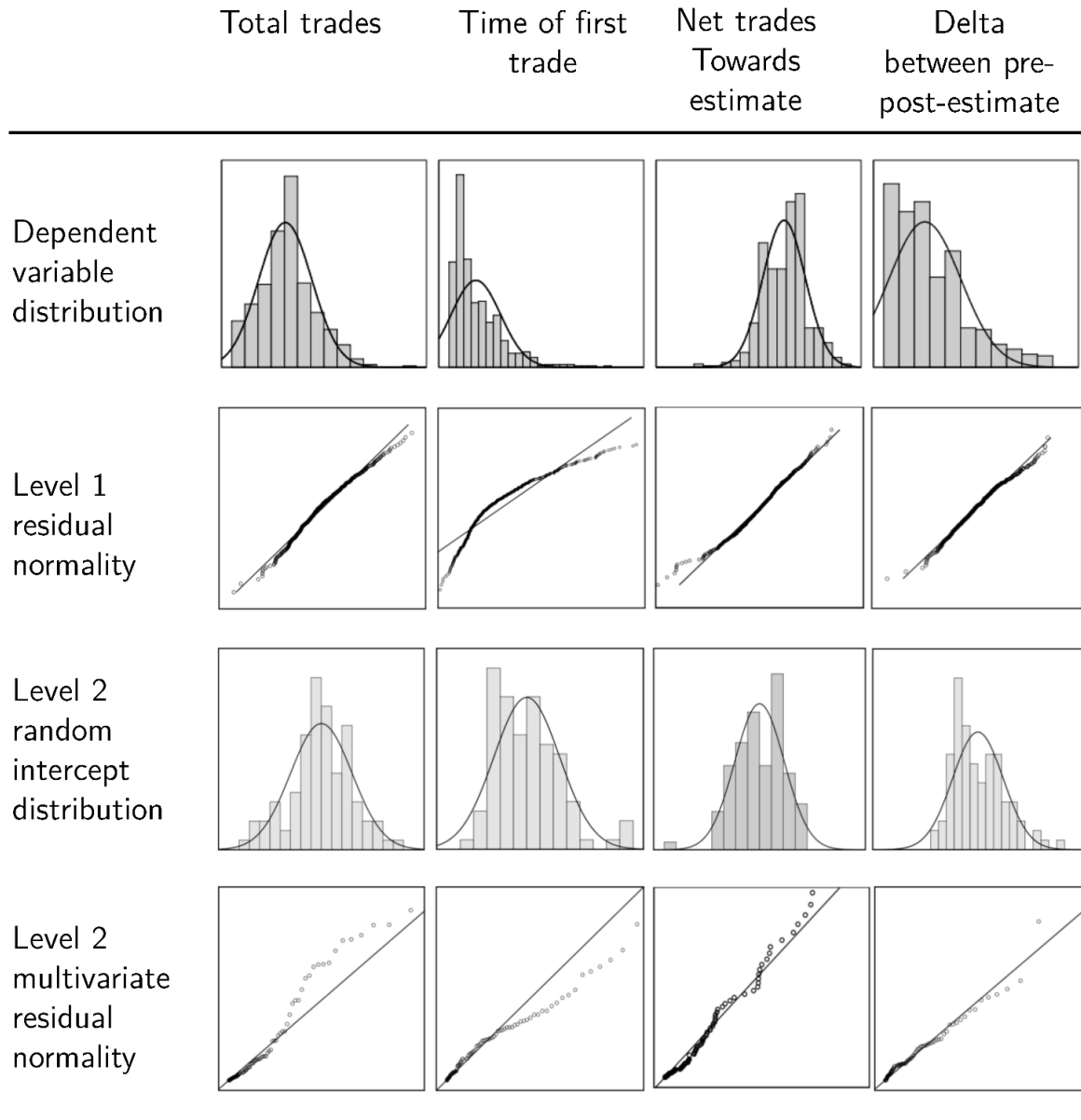


Figure 7.4.: Distributions and residual normality checks for the HLM analysis

pointed out that close normality fit of dependent variables is mainly important for small sample sizes ($n < 50$) and that residual normality is significantly more important for model fit (Hair et al. 2010), deviations should nonetheless be thoroughly investigated. As suggested by Raudenbush and Byrk (2002), we transformed the two corresponding dependent variables via their square roots to fit normal distributions. Afterwards, we built the model with the transformed the dependent variables. The analysis of both models revealed identical coefficient estimates and standard deviations. Additionally, we ran all analyses simultaneously and estimated the model using robust standard errors. Robust variance estimators supply a basis for hypothesis tests, even if distribution assumptions of the HLM are incorrect (Raudenbush and Byrk 2002, p. 278). However, we found insignificant differences between mode-based and robust standard errors, which indicates valid model specification. Accordingly, we decided to hold on to the untransformed variables and model-based estimators, as they can be interpreted more feasibly. Investigating the residual distribution at Level 1 and Level 2, we find sufficient fit overall with the normality assumption. Deviations appear at the edges, where extreme values are situated. The only meaningful deviation from normality can be observed at the “time of first trade” variable. This, however, disappeared when the abovementioned transformations were applied. It was therefore not regarded as a barrier to building the model and investigating its results.

Heteroscedasticity in an HLM violates the assumption of homogeneous error variances at Level 1. Our primary analysis revealed heteroscedasticity for all models. Yet, HLM allows us to account for heteroscedasticity by explaining variance deviations via systematic differences at Level 2. We indeed observe that heteroscedasticity disappears if we explain variance deviations by the treatment condition. This appears sensible because our treatment apparently reduces within-group variance in confidence levels and - in line with our hypothesis - subsequent behaviors.

7.4. Results

As mentioned, the sample covered trading in 521 information market periods in total, each dealing with one innovation evaluation question from the German car market. 7,929 overall trading actions resulted in an average information market stock turnover of 406.1 stocks per subject and market. The results are presented in the same order in which the hypotheses were presented in Section 7.1. All four hypotheses were tested via HLM analysis using the software package HLM 7. Following suggestions by Snijders and Bosker

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

(2012, pp. 102-107), the statistical model was built by starting with an estimation of the fixed effects at Level 1 (Model 1) and then extended by introducing subject-based random effects such as treatment and covariates at Level 2 (Model 2). This section will focus on presenting the hypothesis-related results and will highlight some additional notable findings. The discussion of these results will be carried out in the following section.

The **first hypothesis** claimed that confidence would negatively impact the time until

Variable	Time of first trade	
	Model 1	Model 2
Fixed effects		
<i>Intercept</i>	14.80*** (0.89)	15.30** (0.96)
Trading Distance	-1.07 (1.09)	-1.12 (1.09)
Artificial Agent Trade Direction	1.62** (0.55)	1.57* (0.56)
Random effects		
Treatment condition	-	-3.14*** (0.78)
Gender	-	-0.99 (0.90)
Financial Risk Attitude	-	-0.38 (0.77)
Product Domain Involvement	-	0.24 (0.57)
Deviance	4096.25	4080.46
Deviance Change (d.f.)		15.79 (4)***
Additional Variance explained		4.6%

Notes. n for market level = 521; n for subject level = 99.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.1$.

Table 7.6.: The impact of confidence on the time of first trade (in s)

subjects become active traders during the market periods. The results of the HLM analysis regarding the second hypothesis are presented in Table 7.6. We first looked at the fixed-effects Model 1 before focusing the actual hypothesis test. Model 1 illustrates the results of the fixed-effect covariates, Trading Distance TD and Artificial-Agent Trade Direction AATD. We find that TD does not appear to have a relationship with the first trading time, while the AATD relates significantly to the first trading time. Whenever

the market traded upwards, subjects waited longer to make their first trade ($p < 0.01$). Subjects should see equal opportunities for generating profits in instances where prices moved upwards or downwards away from their initial estimates. However, they took roughly 1.6 seconds longer to make their first trades when prices moved upwards. Model 2 provides a significant reduction of 4.6% in Level 1 variance in the second analysis and hence, significantly improves model fit ($p < 0.001$). None of the random-effects covariates significantly explain any of the error level reduction. Of the additional variables, only the treatment condition significantly relates to the first trading time. In accordance with the second hypothesis, higher confidence leads to significantly earlier engagement in information-market trading ($p < 0.001$). On average, overconfident individuals will trade $2 \times 3.14 = 6.28$ seconds earlier than subjects in the low-confidence treatment condition. Hence, **Hypothesis 1** is not refuted.

Table 7.7 presents the HLM analysis regarding the impact of confidence on overall stock turnover, thus providing a test for the **second hypothesis**. Model 1 first illustrates the results of the fixed-effect covariates, Trading Distance TD and Artificial-Agent Trade Direction AATD. Both have a significant impact on individual stock turnover. Evidently, a pushier artificial trader induces subjects to also trade more intensely. The more the artificial agent opposes their beliefs, the more opportunities arise for subjects to harness profits based on their initial evaluations. When TD was high, subjects traded 46.09 stocks more than in a low TD environment ($p < 0.001$, Model 2). More interestingly, individual stock turnover was negatively related to AATD. If the artificial agent traded towards a target above the subject's estimate, subjects traded 46.24 stocks less than if the computer traded towards a target below the subject estimate ($p < 0.01$, Model 2). This finding may indicate that subjects are more likely to adopt higher estimations than to alter their estimations downwards, which has previously been observed in the context of wishful thinking (Seybert and Bloomfield 2009). Further support for such a notion may become visible when examining the impact of AATD on the net stock turnover towards initial estimates.

Including the subject-based variables in Model 2 significantly decreases the overall model's mean square prediction error, as indicated by the deviance change and the percentage of additional variance explained. 9.5% of additional Level 1 variance can be explained by the adding the subject level variables ($p < 0.001$). Concerning the first hypothesized relationship, confidence indeed exerts a mildly significant and positive impact on trading activity ($p < 0.1$). Though we are reluctant to refute it we only find limited

Variable	Overall stock turnover	
	Model 1	Model 2
Fixed effects		
<i>Intercept</i>	395.12*** (16.31)	378.62*** (17.62)
Trading Distance	44.72*** (12.22)	46.09*** (12.20)
Artificial Agent Trade Direction	-23.93** (8.41)	-23.12** (8.48)
Random effects		
Treatment condition	-	29.07[†] (16.81)
Gender	-	34.74* (16.85)
Financial Product Risk Attitude	-	36.63* (16.85)
Product Domain Involvement	-	-0.43 (11.93)
Deviance	6739.55	6721.53
Deviance Change (d.f.)		17.47 (4)***
Additional Variance explained		9.5%

Notes. n for market level = 521; n for subject level = 99.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$.

Table 7.7.: The impact of confidence on total stock turnover

support for **Hypothesis 2**. Both gender (i.e. being male) and FPRA have larger effects on the sum of individual stock turnover ($p < 0.05$). Overall, stock turnover seems to be triggered more strongly by these two subject characteristics than by confidence.

The **final two hypotheses** explore the relationship between confidence and the willingness to update information and behavior if external information deviates from internal estimates. First, the relationship was analyzed within the information market by looking at the net stock turnover towards initial estimates. Then, we examined the relationship between confidence and the change in pre- and post-market evaluations of the underlying prediction targets, to focus on a market-external variable.

The **third hypothesis** proposed that confidence will be positively related to an individual tendency to trade stocks towards initial estimates, i.e. trading is less likely based on

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

updated beliefs, but to hold on to the original estimate. Preceding the hypothesis test, table 7.8 shows that both fixed-effects covariates TD and AATD are positively related to the net turnover ($p < 0.001$). Comparable to the first HLM analysis, a larger TD induces subjects to trade more stocks towards their estimate because they need to invest more effort to align stock prices with their estimations. AATD is also significantly related to the dependent variable, which strengthens the prior assumption that subjects may indeed have preference for *long* positions. Subjects appear to be less likely to oppose positive price movements via selling than to oppose negative price trends via buying information market stocks. After adding the subject-based variables in Model 2, we find

Variable	Net stock turnover towards initial estimate	
	Model 1	Model 2
Fixed effects		
<i>Intercept</i>	222.60*** (18.35)	196.69*** (18.58)
Trading Distance	123.79*** (14.38)	125.40*** (14.52)
Artificial Agent Trade Direction	-39.28*** (9.26)	-38.09*** (9.14)
Random effects		
Treatment condition	-	69.50*** (19.09)
Gender	-	54.24** (16.86)
Financial Risk Attitude	-	13.48 (14.75)
Product Domain Involvement	-	-20.75 (13.04)
Deviance	6849.95	6820.84
Deviance Change (d.f.)		29.11 (4)***
Additional Variance explained		15.5%

Notes. n for market level = 521; n for subject level = 99.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.1$.

Table 7.8.: The impact of confidence on net stock turnover towards initial estimate

a significant positive relationship between confidence and net stock turnover towards the initial estimate, which gives supports the third hypothesis ($p < 0.001$). The more confident subjects are, the more likely they are to hold on to their original estimates when

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

engaging in information-market trading. We therefore cannot refute **Hypothesis 3**. On average, overconfident subjects traded 139 more stocks towards their initial estimate than subjects in the low-confidence condition. Similarly, gender appears to have had a strong and significant impact on how likely subjects were to align their trading with initial estimates ($p < 0.01$). In contrast to the analysis of overall trading activity, FPRA no longer significantly explains variance in the dependent variable. While it can help to explain overall trading activity, it does not explain the negligence to update expectations from external trading activity when engaging in trades during the information-market periods.

We can further drill down the results for **Hypothesis 3**. Figure 7.5 decomposes trading activity towards and against initial estimates and furthermore partitions trading by the time interval of the trading period and treatment condition. In the graphic, the

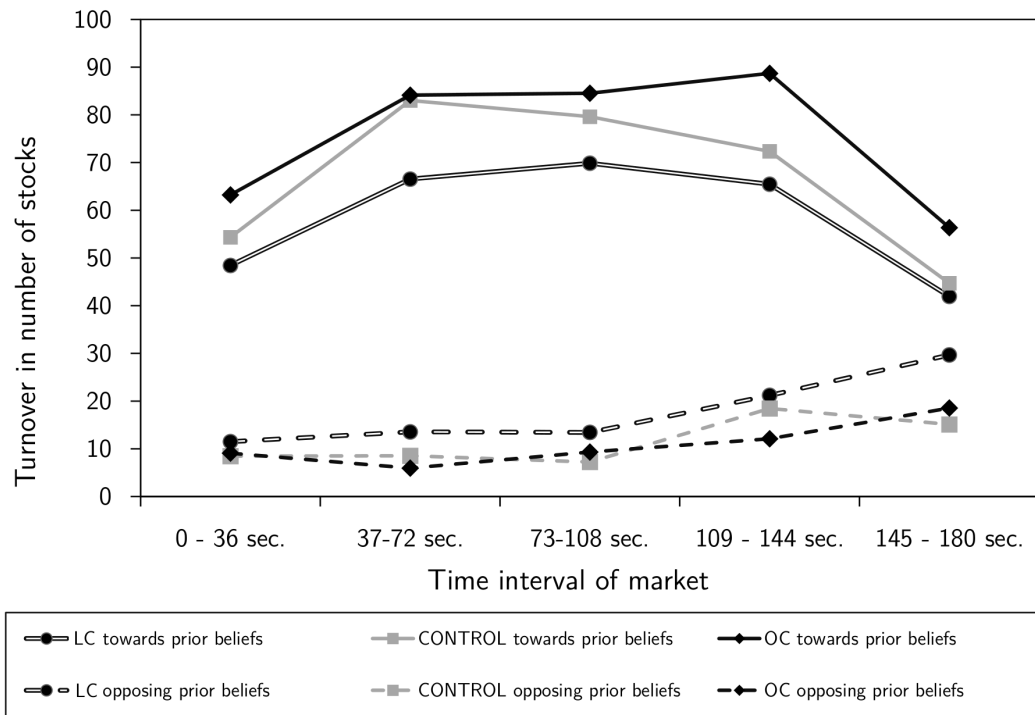


Figure 7.5.: Development of market transactions, by treatment condition, over the market's running time (Source: Own depiction)

overall market running time is split into five segments of equal duration (36 seconds). The ordinate depicts the amount of stocks traded by the treatment group during the time segment, either towards or opposing the initial estimation. We observe that higher

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

confidence indicates significantly higher trading towards initial estimates at all time segments. In addition to that, trading towards initial estimate peaked in the middle of the market's running time. Interestingly, however, trades that opposed the initial estimate paint a slightly different picture. Here, trading only took off at the end a market's running time. Only then did stronger discrimination between the three treatment conditions becomes visible, most notably through the significant increase in opposition trading by low-confidence subjects.

The **fourth hypothesis** focused on the relationship between confidence and the amount

Variable	Delta between pre- and post-market estimate	
	Model 1	Model 2
Fixed effects		
<i>Intercept</i>	8.57*** (0.58)	9.10*** (0.67)
Trading Distance	9.81*** (0.96)	9.78*** (0.96)
Artificial Agent Trade Direction	0.28 (0.35)	-0.28 (0.35)
Random effects		
Treatment condition	-	-2.28* (0.91)
Gender	-	-1.05 (0.83)
Financial Risk Attitude	-	-0.80 (0.72)
Product Domain Involvement	-	-0.84 (0.83)
Deviance	3771.92	3761.24
Deviance Change (d.f.)		10.67 (4)***
Additional Variance explained		5.5%

Notes. n for market level = 521; n for subject level = 99.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.1$.

Table 7.9.: The impact of confidence on the delta between pre- and post-market evaluation

of change from pre- to post-market evaluations of the underlying prediction target. The hypothesis was, again, tested via HLM analysis. The results are displayed in Table

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

Dependent variable	Treatment Condition			Hypothesis
	LC	CONTROL	OC	
Time of first trade in seconds	18.0 s (15.6)	14.1 s (13.6)	11.5 s (12.6)	H1 supported
Overall stock turnover	382.8 (196.9)	391.8 (218.9)	444.6 (182.0)	H2 weakly supported
Net trades towards initial estimate	201.2 (246.4)	276.2 (212.9)	349.6 (212.7)	H3 supported
Deviation between pre- and post-estimate	15.1 (10.6)	14.7 (12.9)	10.2 (10.9)	H4 supported

Notes. n for market level = 521; n for subject level = 99; sd in brackets.

Table 7.10.: Average outcomes for dependent variable by treatment condition

7.9. The TD did have a strong positive effect on the change in subject's estimation ($p < 0.001$). When the TD was far, subjects adjusted their estimation, on average, 9.78 points more than in cases where the TD was small (Model 2). Interestingly, while the AATD had a significant impact on subjective in-market behavior, it did not significantly impact estimation updating.

Concerning the hypothesized relationship between treatment and estimation updating, it can indeed be observed that higher confidence likely leads subjects to change their initial updates less strongly. On average, overconfident subjects changed their pre-market estimations by 4.56 points less than subjects in the low-confidence condition. The negative relationship between confidence and evaluation delta is significant ($p < 0.05$). Our model thus provides support for **Hypothesis 4**.

7.5. Discussion

The experimental study examined the specific behavior related to individual overconfidence when evaluating the future success of innovations via information markets. The results indicate that confidence levels can have a significant impact on individual participation behavior in such group evaluation mechanisms. Table 7.10 summarizes the results and presents the average outcomes by treatment condition. Consistent with **Hypothesis 1**, overconfident individuals are among the first to communicate their predictions regarding the success of innovative technologies and new products compared to less confident peers. This implies that early information cues transmitted via the market mechanism are likely to originate from the evaluations of overconfident individuals.

7. Experiment 1: Overconfidence and Individual Behavior in Information Markets

Given that participants also update their beliefs based on market signals, these early signals and consequently, the beliefs of overconfident participants, may have a relatively strong impact on aggregated beliefs. According to our theoretical basis, the earlier the subjects receive signals, the higher the likelihood that the signal will be processed as relevant information. However, overconfident individuals not only communicate their evaluations earlier. Over the total evaluation period, they are more active, trading significantly (though weakly significantly) more information market stocks, which is consistent with **Hypothesis 2**. Hence, a disproportionately high amount of overconfident traders' actions will be found among all stock price changes. Overconfident participants distribute more private evaluations via the information market mechanism than well-calibrated subjects. Thus, their assessments of innovation success will have relatively big impact on aggregated predictions.

Furthermore, we examined whether overconfidence reduces willingness to change beliefs when contradictory information emerges. Again, our findings are consistent with the hypothesis. We find support for **Hypothesis 3** when looking at market trading. However, we also presented arguments that unobserved trading motives or noise trading very likely hinder a clearer view of the effect of contradictory information on trading via belief updating. It would be highly interesting to observe the behavior of overconfident individuals in field applications where information markets are run over longer periods of time and contradictory information is not only facilitated via market prices but also directly from external sources such as competitors, customers or colleagues. It may be possible that an increase in running time and externally-induced contradictory information will emphasize the positive relationship between overconfidence and unwillingness to update evaluations. Nonetheless, when examining pre- and post-market evaluations in our study, we indeed find highly significant differences in the absorption of market signals to form updated predictions, which is consistent with **Hypothesis 4**. The more overconfident an individual is, the less likely he will be to take market information into account when making a final evaluation after the group mechanism has finished. This is especially important when considering that overconfidence is likely to concentrate among corporate decision makers. Information markets provide them with valuable insights on innovation potential by accessing various sources of previously untapped information. Alas, deciders who should be using the aggregated predictions for making valid evaluations may be most reluctant to do so.

This experiment gives the indication that initiators of group methods to evaluate innovations may need to pay close attention to the participants involved. Evaluation results

7. *Experiment 1: Overconfidence and Individual Behavior in Information Markets*

may incorporate predictions by overconfident individuals disproportionately. This may prove challenging to address, as relevant information may be hidden among individuals who rarely participate in innovation evaluations or other management decisions. They are likely to exhibit less confidence in their evaluations and consequently hesitate to communicate their beliefs via information market stock trading. Relevant information may thus get lost if participants lack confidence. However, such effects may be even more prominent in evaluation methods, where the contribution of valid information is not systematically incentivized. While being prone to overconfidence biases, information markets may still outperform classical forms of information aggregation with regards to the impact of biased participants due to the financial incentives for revealing true beliefs. Here, more understanding about the impact of overconfidence on competing mechanisms such as Delphi is needed.

Our research highlights promising avenues for future research. Most importantly, we did not address whether overconfidence actually improves (or worsens) overall innovation evaluations via information markets. We can infer from our findings that the quality of information market predictions will be highly dependent on the relative quality of overconfident participants' information and their relative presence as participants in the markets. If overconfident individuals hold more valid expectations of innovations' success potential, they will likely improve predictive quality. However, as stated by the literature, overconfidence is often accompanied by overly optimistic expectations, which strongly puts into question the superiority of overconfident individuals' predictions (Lovallo and Kahneman 2003; Trevelyan 2008; Dushnitsky 2009). Further research is required to address the impact of information distribution among participants. Before that is done, evidence that overconfidence leads to different evaluation behavior must not be confused with the implication that the presence of overconfidence generally leads to inefficient innovation evaluations. To address this question, the prediction quality of information markets for innovation evaluation will be explored in **Experiment 2**.

Finally, another important aspect for further investigation relates to the specific environment that was presented to the subjects. In our experiment, participants were not confronted with any information, besides market prices, that could violate their private beliefs. After learning about their potentially superior (or inferior) evaluation skills, they received only market signals. Yet, in real-world settings, subjects will very likely learn about the validity of their evaluations via a more direct means of communication and further information sources.

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

The results of the first experiment indicated that overconfident participants can exhibit specific behaviors when trading in information markets and extracting information from those markets. At least under the given experimental conditions, higher confidence led participants to more aggressive trading activity and less sensitivity to market signals. However, the first experiment did not address the relationship between participants' confidence, the information environment, and the information market's predictive quality. This question is therefore the focus of our second experiment.

The experimental setup used to study the prediction quality of information markets was briefly introduced in Section 5.4. The dependent variable is the **prediction quality of information markets**, with the goal of explaining the relationship between the confidence level of traders and the prediction quality of the information market.

In our experiment, an experimental case consists of one information market period. During each period, one information market stock represents an innovation evaluation outcome. In the market periods, two informed traders, both of whom receive either overconfidence or low-confidence treatment, are joined by two uninformed and untreated noise traders. Accordingly, cases are allocated to a single treatment condition, in which either overconfident traders or low-confident traders are present.

In the basic experimental setting, all treated subjects received the same three pieces of **diagnostic information**. As a result, the treated subjects can be considered well-informed because they have received diagnostic external information, whereas the noise traders can only draw from their personal information bank and market signals to form expectations. This setting is later extended by a scenario in which treated subjects are given the option of purchasing zero, one, or three pieces of diagnostic information at a cost.

8.1. Development of hypotheses

The second experiment focuses on prediction quality as the central dependent variable. In it, we are studying the impact of overconfidence treatment on the prediction quality of information markets.

We consider two relevant prediction dimensions that are influenced by the information market: first and foremost, the information market yields aggregate **market predictions** via the market prices that result from trading; and secondly, the information market and its outcomes may impact **individual predictions** after the markets have finished. The research framework highlights the fact that decision makers often participate in information markets and may (at least to some extent) learn from the results when rendering decisions afterwards. We therefore find it similarly important to also address individual predictions after the information markets have finished in order to test assumptions about how market prices impact individual post-market predictions. Hence, the development of hypotheses branches out into aggregate market predictions and individual post-market predictions.

8.1.1. Overconfidence and prediction quality of information markets

In the case that information aggregation in a market strictly follows the efficient market hypothesis, the market prices' prediction quality will only be subject to the quality of individual information, as it will be efficiently revealed in the trading process.

Yet, Section 4.4 has highlighted research results from the domain of behavioral economics showing that activities and outcomes in experimental and real-world markets often diverge from the predictions of the efficient-market hypothesis. Among other outcomes, speculative trading and information cascades may increase market error when agents do not update or reveal information rationally via trading (Glaser et al. 2003).

The first experiment revealed that individual biases can significantly influence trading behavior, independent of individual information quality. Weber (2000) highlights that participants' information quality and trading behavior needs to be synchronously considered as a central subject-based driver of market prices' predictive quality. We therefore focused the development of our hypotheses on the influence of information quality, trading behavior and the interaction of these two drivers on market prediction quality.

The argument that **more informed traders increase market prediction quality**

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

is economically straightforward. Better information quality among market participants leads to more informed trades, which reduces market price error (Fama 1970). Except for the theoretical case where perfect information among all traders prevents any trading activity (Milgrom and Stokey 1982), the presence of more information will generate better market predictions.

In our experiment, we control the manipulated subjects' information by providing the treated subjects with identical diagnostic cues regarding the information market prediction tasks. Previous research has argued that overconfident individuals will be less likely to absorb diagnostic cues if those cues diverge from personal expectations (Hayward et al. 2006). In our case, it is thought that this can induce greater variance and error in individual predictions before the market by overconfident subjects. Yet, the experimental setup presents the diagnostic cues jointly with the tasks, which supports hindsight among treated subjects. As they are not required to express task-related expectations before the market starts, all treated subjects are expected to form beliefs as if the newly-acquired information had been previously possessed (Astebro et al. 2007; Cassar and Craig 2009). As a result, we can expect their individual predictions before the market to be relatively homogeneous, with variance lower than noise traders' individual predictions and independent of the treatment condition. Consequently, the quality of treated subjects' predictions is not expected to increment average market prediction errors but rather keep them within the boundaries of their informational background.

Noise traders, on the other hand, do not receive diagnostic cues, which would reduce individual prediction errors and variance in prediction error among those participants. Noise traders are thus expected to exhibit highly uninformed individual predictions with large variances in prediction quality. Additionally, the noise traders do not have access to diagnostic information during the market, except for what they can extract from market signals. The information market prices are thus expected to suffer in prediction quality if these uninformed individuals reveal their information via trading, influencing market prices.

Hypothesis 5 *Noise traders' average individual prediction errors before the market will increase average market-prediction error.*

The first experiment demonstrated that the level of individual confidence can significantly impact individual trading behavior. Trading behavior describes how subjects trade in the market and how they respond to other traders' market signals. In particular, overconfident subjects were more likely to act early in information markets, trade more stocks per transaction, and more strongly oppose market signals. Additionally, the

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

post-market predictions showed that overconfident individuals are less likely to update private predictions based on market signals.

Previous research has explored how and why overconfidence and subsequent trading behavior influences the efficiency of information aggregation in markets.

On the one hand, focus can be placed upon **the direct impact of overconfident traders' behavior** on market prices and their predictive quality. Theoretical work has shown that the presence of overconfident traders increases market price volatility (Benos 1998; Caballé and Sákovics 2003; Odean 1998). Market prices are more likely to fluctuate because overconfident traders trade more aggressively. The label of aggressiveness can be reflected in the strength of market signals sent by overconfident traders, such as the amount of shares they trade per transaction (Benos 1998; Caballé and Sákovics 2003) or their overall trading volumes (Odean 1998). As previous research and the first experiment have shown, overconfident traders overvalue the predictive quality of private information, which creates a stronger reaction to market signals that contradict their beliefs (Glaser et al. 2003). The first experiment demonstrated the increase in aggressiveness via increased trading volumes, earlier trading and stronger opposition to market signals that oppose private information.

However, results from existing research on the direct impact of overconfident traders' behavior on the predictive quality of market prices are ambiguous (Glaser et al. 2003). On the one hand, Benos (1998) have provided a model in which all traders are well informed. Overconfident traders increase liquidity and market depth, which leads to quicker convergence and efficient market pricing. On the other hand, Nöth and Weber (2003) have produced experimental evidence showing that overconfident traders' behavior negatively influences market price quality. Even when relatively long sequences of market signals favor outcomes that oppose overconfident subjects' private evaluations, they will too often break such information cascades, thus reducing the aggregate quality of the market signals.

In the context of our experiment, the direct impact of overconfident traders' more aggressive trading behaviors on market prediction quality is ambiguous. They are expected to more strongly oppose ill-informed noise traders because they would be less likely to absorb those signals as valid information, which could increase the amount of correcting market signals. Yet, they would also be more likely to strongly oppose the signals of the other well-informed overconfident traders. This could prevent both treated subjects from learning from the respective interpretations. In summary, it is unclear whether

overconfident traders' behavior will directly impact market prediction quality.

On the other hand, focus can be placed upon **the indirect impact of overconfident traders' behavior** on market prediction quality via the reaction to their signals and subsequent action of other traders.

First, theoretical insight from signal detection theory has helped us understand how aggressive trading behavior by overconfident subjects yields different signal perception and processing by noise traders than less aggressive trading behavior by low-confidence subjects. More aggressive trading increases market price volatility. Experimental evidence based on signal detection models shows that higher volatility in the values of sequential signals decreases subjects' performance in extracting valid information from these signals (Macmillan 2002; Gold et al. 2004). Subjects are less likely to identify underlying commonalities when signal values vary more strongly.

In a context more closely related to information markets, Daniel et al. (1998) built his financial market model with overconfident traders on the assumption that overconfident traders' market signals exhibit a larger proportion of noisiness because they do not sufficiently consider previously revealed information.

These theoretical assumptions find support in cognitive and business-related experiments. In a meta-analysis of 13 social learning games, Weizsäcker (2010) shows that lack of agreement in other participants' signals makes it more difficult for subjects to make correct inferences from these signals, thus negatively influencing decision quality. Kremer et al. (2011) find support for the system neglect hypothesis in a forecasting experiment by Massey and Wu (2005), according to whom, individuals place too much emphasis to signals they receive relative to the system that generates the signals. Translated to the information market, subjects may pay too little attention to the information that underlies the overconfident subjects' trading behavior.

Noise traders may furthermore emulate treated individuals' trading behavior. It has frequently been demonstrated in social game experiments that the behavioral differences of manipulated subjects are reflected in the behavior of untreated subjects in interactive scenarios (Camerer and Fehr 2006). This has mainly been attributed to social interaction concepts such as reciprocity and conditional cooperation (Engel et al. 2011). Subjects are sensitive to other group members' behavior and tend to mimic it, independent of potential consequences.

In the context of market behavior, this can augment noise traders' inability to learn from market signals in the case that overconfident traders are present. Noise traders

will be less likely to absorb overconfident subjects superior information before engaging in trading if they mimic their aggressiveness (Bikhchandani et al. 1998).

To summarize, overconfident participants' trading behavior induces noise trader reactions that positively impact their influence on market errors. Such a relationship hence augments the positive relationship between noise traders' individual prediction errors and the information market prediction quality.

Hypothesis 6 *The presence of overconfident traders will increase the positive effect of noise traders' average individual prediction errors before the market on overall market prediction error.*

Individual post-market predictions

In our experiment, post-market predictions are submitted directly after the markets have finished. To make these predictions, treated subjects draw from the diagnostic information received previously, their subsequent pre-market predictions, and the signals they have gathered during trading.

Treated subjects engage in a costly cognitive task to analyze the provided pieces of information and transform them into private estimations. The information provided and subsequent private predictions are represented by a very limited set of four distinct values during this task. The three values provided as diagnostic information are externally labeled as valid information. Treated subjects process and aggregate these values privately to form private pre-market evaluations.

In contrast, subjects do not perceive market signals as similarly valid sources of information. First, subjects have no information about the processing capability of the other participants. Previous empirical work has shown that participants' information market trading is impacted significantly more by predictions from direct and deterministic information sources, e.g. from colleagues in geographic proximity who give specific suggestions and provide specific values, than by market prices (Cowgill et al. 2008). Lack of information about others' capabilities leads to the underestimation of those capabilities (Camerer and Lovo 1999). In a related experiment, Radzevick and Moore (2011) demonstrated experimentally that subjects are more likely follow unambiguous suggestions. Second, the market will provide signals that do not reflect the provided information because noise traders cannot draw from that information. This will decrease the likelihood that the treated subjects will update private estimations based on market signals, as discussed in the market-related hypothesis. They would have reason to doubt the validity of market signals in the light of their private information.

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

The discrepancy in the perceived quality of signals before and during the market emphasizes the relative importance of private pre-market predictions in the post-market predictions of treated subjects. Market signals will have less relevance for forming post-market predictions for treated subjects. Furthermore, overconfident subjects are even more likely to draw inferences for post-market predictions from their pre-market predictions (Deaves et al. 2009). Underlined by the results of the first experiment (see Section 7.4), overconfident subjects are less likely to update their beliefs than subjects in the low-confidence condition. This may increase the positive influence of pre-market errors on post-market errors for overconfident traders. Overconfident subjects are more likely to maintain particularly high pre-market errors because they do not consider market signals, even in the case that they would prove helpful.

Hypothesis 7 *Treated subjects' post-market prediction errors will be positively correlated with their pre-market prediction errors.*

Hypothesis 8 *Overconfidence treatment will increase the positive relationship between individual pre-market error and individual post-market error in treated subjects.*

In our experiment, noise traders did not receive diagnostic information before the market. As a consequence, they were expected to perceive their private estimations as both less valid overall and less valid relative to market signals and compared to treated subjects. This would increase their absolute and relative sensitivity to market signals for use in forming predictions, as discussed with the market-related hypothesis. They would be more likely to update beliefs because, compared to treated subjects, noise traders would be less inclined to engage in biased hypothesis testing based on previously processed information that they ought to deem valid. They would not possess diagnostic information that could narrow the interval of plausible private evaluations or prevent them from dismissing market signals as uninformative.

Two consequences would be expected to arise from this. First, the improvement of noise traders' predictions between markets would depend primarily on their pre-market prediction error. The higher their pre-market errors, the more likely they would be to improve predictions after updating beliefs because the market would have provided them with signals to allow for prediction improvement. They would be able to harness the superior information of the treated subjects to improve their own private predictions. Second, the market error would significantly influence the quality of their post-market evaluations. As they would indeed draw information from market prices to update their

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

beliefs, higher market errors would limit noise traders' ability to improve their predictions from pre-market errors.

Hypothesis 9 *Noise traders' post-market prediction errors will be positively correlated with their pre-market errors.*

Hypothesis 10 *Noise traders' post-market prediction errors will be positively correlated with overall market error.*

8.2. Experimental design

8.2.1. Participants

We recruited 136 graduate engineering students from the same university as in the first experiment to participate in the information markets. The students were part of the consecutive class that followed a year after the participants of the first laboratory experiment. This ensured that subjects had similar educational backgrounds and demographic characteristics but no knowledge about the previous experiment. All held bachelors' degrees in engineering disciplines and were currently enrolled in business administration courses at Hamburg University of Technology. The average age was 23 and 20 percent were females. We again issued pre- and post-experimental questionnaires to evaluate financial risk attitude, product-domain involvement, and individual predictions regarding the innovation-evaluation tasks. This was particularly necessary to evaluate the effect of participation on the prediction quality in the innovation evaluation tasks. After the market exercise, we presented each subject with a similar information market quiz to test their understanding of incentives and the information market in general. This time, all 136 subjects participated successfully in the experiment.

8.2.2. Innovation evaluation tasks

The second experiment was aimed at investigating the relationship between overconfidence and prediction error in the context of innovation evaluation. For this, we required **innovation evaluation tasks where the outcome was truly unknown at the time of the prediction** so that the information market prices and individual predictions could be analyzed in terms of predictive quality. Hence, the experiment required the application of tasks that relate closely to innovation evaluation, featured unknown future outcomes at the time of prediction, and still allowed for us to provide diagnostic information. Similar to the first experiment, the evaluation tasks were related to success

potential and marketing-related characteristics of innovation products, such as market share developments of innovative products and product characteristics.

We partnered with **GfK Retail & Technology GmbH**, the global leader in tracking the sales of technology-based consumer goods such as domestic appliances, office technology, consumer electronics, and communication technology, to define relevant prediction tasks for innovative products in the category of mobile computing and derive diagnostic information based on their expertise. At the same time, the company's market data allowed us to evaluate prediction errors after the markets had finished and the prediction events had occurred to test our hypotheses.

We chose the **mobile computing product category** for three reasons. First, that product category is highly dependent on valid innovation evaluation, since product development costs and failure rates can be very high (Eisenhardt and Tabrizi 1995). Second, the subject group had been highly exposed to the innovative consumer electronic products upon which we focused the evaluation tasks, both as frequent users of novel devices but also in the academic context of innovation management and marketing courses. Finally, high-technology companies such as consumer electronics firms are considered a similarly attractive future employer for German engineering students as the automobile industry, with companies such as Google and Apple ranking on a par with leading automobile manufacturers in Germany (Trendence Institut 2011). Thus, our subjects were expected to exhibit a sufficiently high degree of product involvement and were initially assumed to be capable of making informed predictions.

We developed a set of 16 innovation-related questions ranging from the assessment of potential market success for specific tablet-PC products such as the Apple iPad, to evaluating the price drops of new technologies such as SSD hard drives.¹

Analogous to the first experiment, 10 of the 16 tasks were randomly drawn for use in the manipulation treatment to induce different confidence levels via false feedback. The remaining six tasks were then used in the information markets.

The three pieces of additional information per information market period were developed in collaboration with market experts. We ensured the **diagnosticity of the presented information** by consulting with senior GfK market experts and by drawing on the company's internal forecasting systems. Thus, the information can be considered diagnostic from an objective perspective. It was furthermore emphasized while developing the cues that the information pieces should add information quality if presented

¹The complete set of evaluation questions and information market prediction tasks are included in the Appendix.

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

sequentially. After having generated the information market questions and the pieces of information, we conducted an informal pre-test with a non-related lecture group of 50 industrial-engineering master students who had similar characteristics but who would not be participating in the information markets. The pre-test showed that the information given led to better individual predictions and that individual predictions improved with the amount of information received, when using GfK forecasts as proxies for prediction outcomes.

8.2.3. Implementation

The experiment was again executed using z-Tree, a software framework for programming economic experiments (Fischbacher 2007). Figure 8.1 depicts the design of the second

Experimental setting	Overconfidence treatment: 2 subjects are treated with a <u>positively</u> manipulated feedback, 2 subjects do not receive any feedback	Low-confidence treatment: 2 subjects are treated with a <u>negatively</u> manipulated feedback, 2 subjects do not receive any feedback
<u>Basic experimental setting:</u> The 2 treated subjects receive three pieces of diagnostic information at no charge.	9 groups á 4 subjects x 5 information-market periods = 45 market predictions and 180 individual post-market predictions	9 groups á 4 subjects x 5 information-market periods = 45 market predictions and 180 individual post-market predictions
<u>Additional experimental setting:</u> The 2 treated subjects can acquire either 0, 1 or 3 pieces of diagnostic information.	8 groups á 4 subjects x 5 information-market periods = 40 market predictions and 160 individual post-market predictions	8 groups á 4 subjects x 5 information-market periods = 40 market predictions and 160 individual post-market predictions

Figure 8.1.: Design for the second experiment, by experimental setting: Number of cases by treatment and experimental setting

experiment and the expected number of cases per experimental cell, based on the subject sample size and market repetitions per trading group. The figure visualizes the second experiment as two 2 X 1 experiments: first, the basic experimental setting in which diagnostic information is provided free of charge to treated participants, and then the extended scenario, in which the subjects have the opportunity to acquire the information

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

at a cost.

For each setting, **the experiment consisted of five parts** that resembled the experimental procedure of the first experiment depicted in Figure 7.3 of the previous chapter. At the beginning of the experiment, subjects randomly drew their seats in the laboratory. There were eight workstations and four subjects were placed together in one trading group.

First, the subjects filled out pre-experimental questionnaires to collect unbiased measures of individual domain-specific innovativeness, familiarity with financial markets, and financial-product risk aversion. They then watched an introductory video that familiarized them with the experimental process and described how to participate in the information market.² The **second part** consisted of the initial evaluation task. Subjects were presented with 10 initial evaluation questions. Again, they were told that their chance to win a €50,- voucher would be influenced by how many questions they answered correctly and how well they estimated their own performance. We kept the incentive for the initial evaluation task in order to obtain an individual baseline measurement for natural confidence. The more confident an individual, the higher the distance between truly correct answers and estimated correct answers will be, given the incentive above.

The manipulation followed in the **third part** of the experiment. Two of the four subjects were randomly selected to jointly receive either the positive or the negative feedback, while the other two subjects entered the information market without treatment, as noise traders. These subjects did not receive any feedback regarding their performance in the evaluation task, nor were they given the opportunity to access diagnostic information during the information markets.

The information market task marked the **fourth part** of the experiment and again consisted of six independent market periods, the first of which was considered a training period. Hence, each group produced five market predictions and ten individual pre- and post-market estimations.

Before each market period, subjects in the treatment condition underwent the information acquisition phase. If the session's scenario indicated the **basic experimental setting with free information**, subjects who received a treatment were given three pieces of diagnostic information regarding the prediction object. If the session's scenario was **additional experimental setting where information had to be acquired at**

²The version of the video for treated subjects in the basic experimental setting may be viewed here: <http://youtu.be/zv8s3zeV8L4>

a cost, subjects who received a treatment could choose between acquiring no additional information, one piece of diagnostic information for two a Euro deduction from a potential 20,- Euro in winnings, or three pieces of diagnostic information for a four Euro deduction from a potential 20,- Euro in winnings.

All subjects were asked to report estimates of the true values before and after trading to measure individual prediction errors while considering information acquisition and information market trading. Each information market ran for 140 seconds. Different from the first experiment, the markets' starting prices were always set to 50 so that the subjects would not strategically set starting pricing to misguide other traders. Prediction tasks and periods were arranged over sessions so that their potential impact could be controlled for in subsequent analysis.

We facilitated trading through the same automated market maker using Market Scoring Rules to update prices, as in the first experiment. Figure 8.2 shows the trading interfaces for subjects in both treatment conditions and the free information condition. The top left corner shows the treatment manipulation and the lower left corner the current prediction task. The freely provided information is depicted in the lower left corner. The stock price development is shown in the top right corner and participants could buy or sell either big (50) or small (5) packages of stock in the lower right corner, which also displays current budget and portfolio. After all markets had closed, two prediction tasks were selected randomly and the subjects were paid out based on the true underlying value of their evaluations. The final rank determined the probability of winning a prize. Accordingly, the subjects' chances of winning two vouchers worth 20 Euros minus any information acquisition costs were determined by their overall portfolio value in these two markets.

As before, the **last part** of the experiment consisted of the post-experimental questionnaire and was very similar to the one used in the first experiment. We first provided questions to evaluate subjective performance perceptions and subjective reasons for the respective performances. The questionnaire then asked subjects about how much they enjoyed the experimental tasks. Finally, the questionnaire addressed whether the subjects understood the information market correctly and collected information about the subjects' age, gender, place of origin, and course of study.

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets



Figure 8.2.: Information market interaction interface for subjects in the overconfidence (top) and low confidence treatment (bottom) groups, in the free information condition

8.2.4. Instructions

The day prior to the session, each subject received a written document with detailed instructions regarding all aspects of the innovation evaluation tasks (except the treatments), the incentives and how to trade and compete in the information market.

Based on the results from the first experiment, the instructions were extended using scenarios of potential trading situations to further foster market understanding. We re-edited and updated the tutorial video to account for the information conditions for subjects in the treatment conditions. This time, three specific versions of the video were produced to fit the information environment of treated subjects in both experimental conditions and the noise traders.

8.3. Data preparation

8.3.1. Variable operationalization and construct validity

Table 8.1 presents the coding scheme for the dependent and independent variables that were included in the statistical analyses. The upper part of the table shows the dependent variables. We chose the root square error (RSE) to measure prediction error for individual and information market predictions because the RSE is a widely accepted measure of the difference between predicted observed values (Armstrong and Collopy 1992). Given that we ultimately obtained a true outcome for each prediction i , we could calculate for each prediction i :

$$RSE_i = \sqrt{(Prediction_i - True Outcome_i)^2}$$

We additionally wanted to analyze the impact of our treatment on information buying in the costly information scenario. We included one measure for information purchases by individual traders and one measure for purchases per trading group.

The independent experimental variables on the individual level did not require any preparation besides the coding scheme displayed in the lower part of Table 8.1. Similar to the first experiment, independent control variables for gender, FPRA and PDI were included in the analysis. The individual estimation errors of the two treated and untreated individuals entered the market-based analysis as group averages. While such a procedure did not allow the accounting of within-group variance in individual predictions, it was considered the only valid option for measuring the impact of individual variables on group

Variable	Operationalization
Dependent variables	
Individual pre-market error	RSE of an individual estimate before an information market period
Individual information Market error	Sum of information acquired per subject RSE of an information-market estimate based on average market prices per market period during the last 20 trading seconds
Individual post-market error	RSE of an individual estimate after an information-market period and market period
Noise traders' post-market prediction improvement	Noise traders' pre-market error – noise traders' post-market error
Independent variables	
Treatment condition	-1 (Low confidence) 1 (Overconfidence)
Gender	-1 (Female) 1 (Male)
Financial Product Risk Attitude	Metric coefficient (<i>extracted via regression from factor analysis</i>)
Product Domain Involvement	Metric coefficient (<i>extracted via regression from factor analysis</i>)
Average Group Error Treatment	Average individual pre-market RSE per market of the two treated traders per market period
Average Group Error Noise	Average individual pre-market RSE of the two untreated traders per market period
Market information	Sum of information acquired per market period

Table 8.1.: Operationalization of variables included in the statistical analysis for Experiment 2

Items	Product Domain Involvement	Financial Products Risk Attitude
PDI1	0,827	-0,007
PDI2	0,886	0,011
PDI3	0,742	0,253
FPRA1	0,026	0,771
FPRA2	0,065	0,791
FPRA3	0,116	0,742
Variance explained	0,34	0,31
Cronbach's α	0,76	0,66

Notes. $n = 128$; Exploratory Factor Analysis: Principal Component Analysis (Eigenvalues > 1); Varimax-Rotation.

Table 8.2.: Measurement of constructs for control variables

outcomes via multi-level regression (Van Kleef et al. 2010). However, individual-variable aggregates needed to be introduced with caution and to a limited degree because they likely to reduce the explanatory power of the statistical model. Again, the two construct-based control variables were investigated and prepared via factor analysis. The factor analysis revealed that all relevant factor loadings are > 0.6 . As can be also seen in Table 8.2, cross loadings are again very low. The PDI construct explains 34% of total item variance, whereas the FPRA construct explains 27% of total item variance. Finally, the factor values for each subject were again extracted via multiple-linear regression.

8.3.2. Data adequacy and statistical model

The second experiment also used a within-subject design with multiple measurements per subject, which again required multi-level analysis. The adequacy requirements that were presented in Section 7.3.3 remain unchanged. A distinction needed to be made between dependent variables at the subject level and dependent variables at the market level. When focusing on a subject-level dependent variable, we used subject-level variables as Level 2 grouping variables. When focusing a market-level dependent variable (e.g. market error), we used market-level variables as Level 2 grouping variables. As we considered two independent information settings (free and costly information) at two different analysis levels (subject and market), we faced a relatively large set of descriptive statistics and tests for data adequacy with HLM regression. Hence, this section will more narrowly focus the discussion on model fit, deviations, and the measures taken to

overcome those deviations.³

First focusing on the **basic experimental setting**, we find deviations from the normality assumptions for the dependent variables at both the individual and group levels. Since the treated subjects received diagnostic information, we find that prediction errors of treated subjects are clustered around the low end of errors for individuals and groups. Accordingly, the distribution of estimations is potentially skewed to the left (i.e. to the lower range of prediction errors). Indeed, dependent variables for individual and market estimates show significant skewness above the often-cited upper bound of two times the skewness' standard error (Miles and Shevlin 2001). We approached the issue by following the procedures previously applied in Section 7.3.3. We ultimately relied on the use of robust variance estimators if the dependent variables were not distributed normally (Raudenbush and Byrk 2002, p. 278). Residual distributions provide sufficient fit with the normality assumption at both levels of analysis after graphical analysis and thus do not prevent statistical-model validity.

Regarding the **additional setting with costly information**, we first needed to consider that regressing the amount of acquired information requires the application of ordinal regression, as three ordinal categories characterize information buying as a dependent variable. Here, we could disregard the dependent variable's distribution and find assumption fit for the remaining criteria. Concerning the other variables, we found similar skewness for individual and market estimations and again, applied the proposed data checks that were previously introduced in Section 7.3.3 (Snijders and Bosker 2012, pp. 153-173). We find sufficient model fit overall after introducing the random effect variables at Level 2, again relying on robust standard errors for estimation. Graphical analysis shows that residual variance is sufficiently aligned with the underlying normality assumption for HLM. We therefore build the models in a straightforward manner.

8.4. Results

In total, 136 subjects participated in the second experiment. This included 34 trading groups, each of which traded in five information markets, thus generating 170 information market predictions. In the following sections, the results will be presented according to the development of hypotheses. We first focus on the impact of overconfidence on the prediction quality of information markets in the basic experimental setting. We will present the market-related results and then move on to individual post-market results.

³All corresponding graphical analyses can be found in the Appendix

Afterwards, we will continue with the results of the additional experimental setting and show that they are fundamentally consistent with an environment where individuals need to acquire diagnostic information.

Comparable to the results from the first experiment, we build the models by estimating fixed level effects at Level 1 (Model 1) and subsequently extend the models by subject- or market-based random effects at Level 2 (Model 2).

8.4.1. Overconfidence and the prediction quality of information markets

We begin by presenting the results for the basic experimental setting in which information was provided to treated individuals free of charge. In total, 72 students participated in the experiment, generating 360 individual predictions before and after the information markets, and 90 market predictions. We will align the presentation of results with the process of the experiment. We begin by presenting the results regarding the market-related hypothesis. Analyzing market behavior in detail will then provides further explanation of the market-related results. Finally, we focus on the hypotheses that relate to the individual predictions after the market were finished.

Market predictions

Table 8.3 shows the result for the first HLM regression. The models draw from ninety cases coming from 18 experimental sessions with 5 periods each. The dependent variable is the market error measured by the average error of market prices during the final 20 seconds of trading.

Model 1 leaves out the treatment condition as a random-effect variable and only regresses the market error on the intercept and the respective average group errors. Model 2 introduces the treatment condition as a market-level independent variable. We added Model 3 to highlight the benefits in model fit when including the intercept market error for the random effects model.

Model 1 shows a significant positive influence of average noise traders' pre-market errors on market error ($p < 0.001$). Yet, the average pre-market errors of treated subjects did not significantly increase average market errors.

These findings are reinforced in Model 2. The model explains 6.1% more variance than the first model, a highly significant effect ($p < 0.001$). Again, average market error was strongly influenced by the average pre-market errors of noise traders ($p < 0.001$)

Variable	Market error		
	Model 1	Model 2	Model 3
Fixed effects			
<i>Intercept</i>	4.79*** (1.99)	6.41*** (2.20)	-
Group Error Treatment	0.06 (0.13)	0.08 (0.14)	0.21* (0.09)
Group Error Noise	0.40*** (0.11)	0.33** (0.10)	0.52*** (0.07)
Random effects			
Treatment condition	-	-1.85 (2.20)	-2.93 (2.79)
Group Error Treatment × Treatment condition	-	0.03 (0.14)	0.05 (0.17)
Group Error Noise × Treatment condition	-	0.31** (0.10)	0.30* (0.11)
Deviance	636.68	604.42	610.57
Deviance Change (d.f.)		32.26 (8)***	6.2 (1)*
Additional Variance explained		6.1%	-3.1%
			vs. Model 2

Notes. n for period level = 90; n for market level = 18.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.1$.

Table 8.3.: Impact of the presence of overconfident individuals on market prediction error in the basic experimental setting

but not so by the average pre-market errors of treated subjects. The direct effect of treatment condition and the interaction effect between treatment condition and treated subjects' pre-market errors are not significant. However, the interaction between treatment condition and noise traders' pre-market errors is significant ($p < 0.01$). Market errors increased 0.62 points more with each incremental increase in noise traders' pre-market error, when the market period featured overconfident subjects.

The last model omits the intercept. The model explains significantly less variance than the second model ($p < 0.05$). We find similar significant relationships as in the second model but without the intercept, the relationship between treated subjects' pre-market errors and market error becomes significant ($p < 0.05$). Still, market errors are much more strongly related to noise traders' pre-market errors in terms of effect size and significance level ($p < 0.001$).

Hypothesis 5 posits a positive influence of noise traders' pre-market errors on average information market errors. As described above, the statistical analysis reveals strong support for this relationship. Even after incorporating random effects in Model 2, we still observe a significant direct effect of average noise trader group-error on market-prediction error. Each point increase in average noise trader prediction error increased market prediction error by 0.3 points ($sd = 0.10$).

Average market error across conditions ($\bar{x} = 12.13$, $sd = 8.79$) was similarly high as average pre-market errors by treated subjects (OC: $\bar{x} = 12.82$, $sd = 11.57$; LC: $\bar{x} = 12.42$, $sd = 11.86$) but significantly lower than noise trader pre-market errors (OC: $\bar{x} = 18.64$, $sd = 13.64$; LC: $\bar{x} = 18.46$, $sd = 11.64$). Furthermore treated subjects had, on average, more homogeneous estimations before the market than noise traders. While the average distance between the two treated subjects' pre-market estimations were $\bar{x} = 7.39$ points ($sd = 7, 33$), noise traders' estimations lay $\bar{x} = 12.39$ points ($sd = 9, 11$) apart before the markets started. A closer relationship between treated subjects' pre-market errors and overall market error, and less variance between treated subjects' private estimations, may explain the finding that only noise traders' pre-market errors significantly altered average market error. Additionally, the treated subjects could reflect market prices in the light of the information provided during the market. Treated subjects more likely reduced the boundaries by which market errors varied around their average.

The statistical analysis further reveals a positive interaction effect between present overconfidence and noise traders' pre-market errors on market error, which concurs with

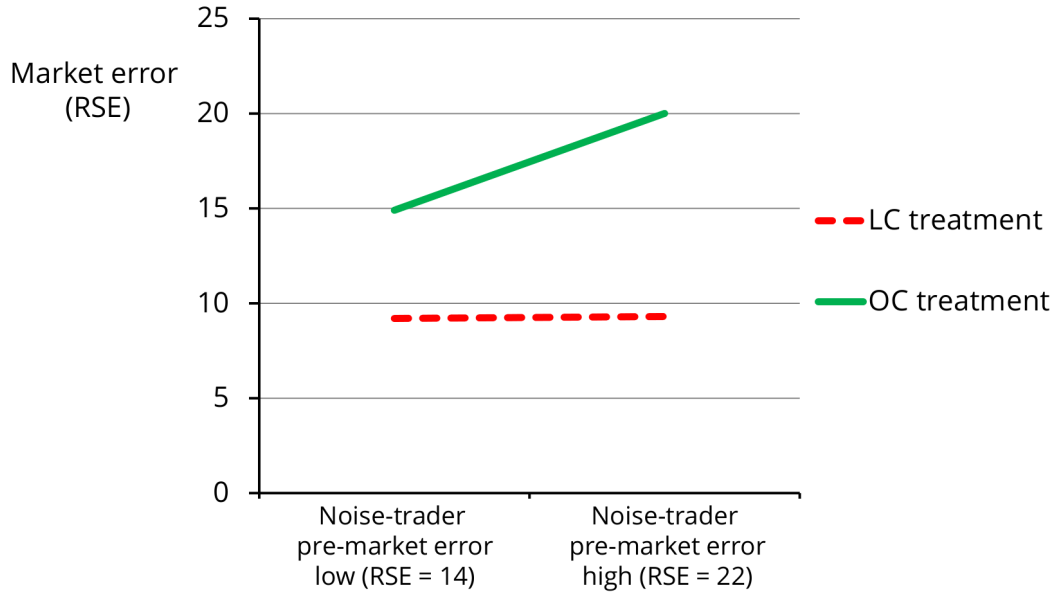


Figure 8.3.: Interaction effect of noise trader pre-market error and the presence of overconfident traders on market error (Source: Own depiction)

Hypothesis 6. When overconfident traders are present, noise traders' pre-market errors will have a more negative influence on the quality of market predictions.

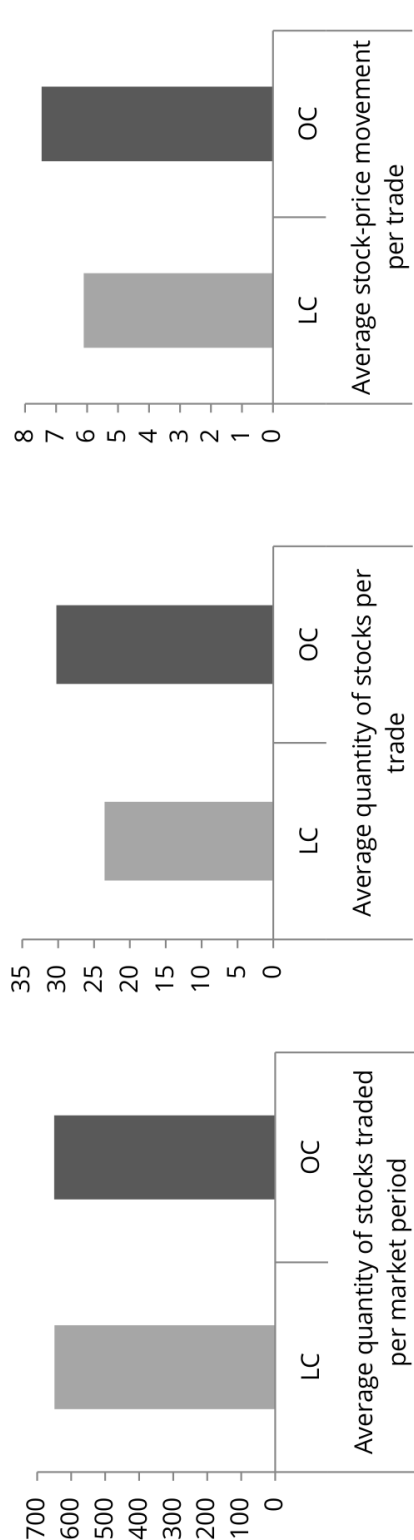
Figure 8.3 documents the interaction effect. Market error was calculated via the regression function from Model 2, which explained most of the variance. First, the average pre-market errors by treated subjects ($\bar{x} = 12.50$) were entered into the function. Then, one below-average ($\bar{x} = 14.00$) and one above-average pre-market error ($\bar{x} = 22.00$) by noise traders was chosen to visualize the interaction effect. The figure shows that an increase in pre-market errors by noise traders had almost no visible effect on market error when subjects with low confidence were present ($\Delta RSE = 0.1$) but significantly increased market errors in the presence of overconfident subjects ($\Delta RSE = 5.1$).

We argued that the interaction effect would be rooted in the particular trading behavior of overconfident subjects, which would furthermore translate into signal absorption and trading behavior of noise-traders.

Consequently, it is sensible to begin additional analysis by more closely inspecting data for indicators of more aggressive trading by overconfident subjects.

Figure 8.4 illustrates relevant indicators of trading behavior by treated subjects. Trading behavior has been analyzed and illustrated separately for overconfident (OC) subjects

a. Trading-behavior indicators for treated subjects



b. Trading-behavior indicators for treated subjects in case trades oppose the previous market signal

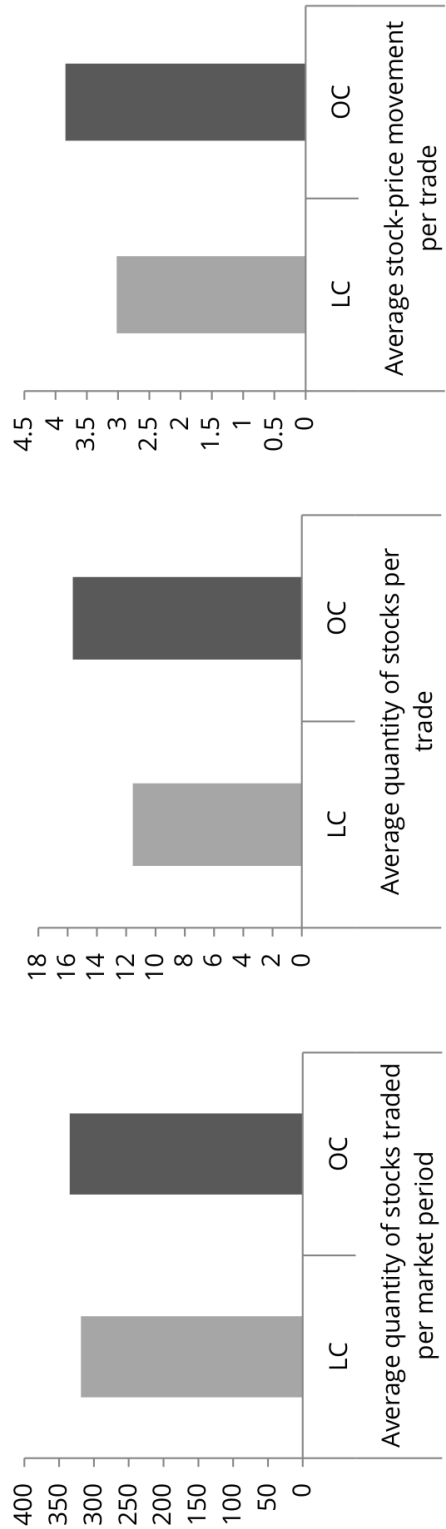


Figure 8.4.: Indicators of trading behavior by treated subjects and treatment condition (Source: Own depiction)

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

and low-confidence (LC) traders. The first row (a) refers to overall trading behavior. The second row (b) only refers to trading behavior concerning trades that opposed the previous trades' market signal. For example, if the previous trade increased the market price, an opposing trade, by our definition, would then decrease the market price. For each row, we selected three variables to indicate trading behavior. The first bars display average sum of stocks traded per market period. The second set of bars show the average amount of stocks per trade by treated subjects. The last set of bars shows the average stock price movement per trade by treated subjects.

We first observe that overall trading volume did not significantly differ between OC and LC subjects. OC subjects traded a few more stocks on average per market period (LC: $\bar{x} = 648,9$, OC: $\bar{x} = 649,9$) and when opposing previous market signals (LC: $\bar{x} = 318,7$, OC: $\bar{x} = 334,7$). This finding concurs with the results of the first experiment. Even when the artificial market agent continuously opposed market signals, OC subjects barely traded more stocks per period ($p < 0.1$, see Section 7.4).

Average quantity of stocks per trade refers to the amount of stocks that were traded in each market section. The participants could trade either 5 or 50 stocks with each market action. Here, we find a different picture and indeed, significant differences ($p < 0.01$, independent sample t-test) between the trading behavior by OC and LC subjects. Over all trades, OC subjects traded 6.7 stocks or 29% more per trade than LC subjects. While average stocks per trade were lower in cases where the treated subjects opposed previous market signals, OC subjects still traded 4.1 more stocks per trade (35%, $p < 0.01$, independent sample t-test) than LC subjects. These differences directly translate into the average stock price movements depicted in the last columns of row (a) and (b) of Figure 8.4. Overall and opposing stock price movements were significantly higher per trade in OC conditions. Each market action by OC subjects can be regarded as more vigorous, as each resulted in larger stock prices changes than actions by LC subjects. Other indicators support the notion of more vigorous trading by OC subjects. While LC subjects committed significantly more trading actions ($p < 0.05$, independent sample t-test) on average, overall, per market (LC: $\bar{x} = 27.6$, OC: $\bar{x} = 21.7$), OC subjects incurred significantly larger "streaks" of market price movement ($p < 0.05$), i.e. consecutive movements of market price by the respective subject (LC: $\bar{x} = 8.5$, OC: $\bar{x} = 9.6$, independent sample t-test).

Our hypothesis implicated that such aggressive and opposing trading behavior would likely prevent noise traders from extracting valid information from market signals. An appropriate proxy for measuring noise traders' signal absorption and learning from mar-

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

ket signals could be to observe how their post-market estimates differ from their pre-market estimates. Assuming that they change estimates for the better, more change should relate to more signal absorption and more learning from the market under the given experimental setting because the treated subjects possessed and provided superior information. On average, noise traders who participated together with LC subjects changed their estimations significantly less ($p < 0.05$, independent sample t-test) than noise traders who joined overconfident subjects (LC: $\bar{x} = 11.1$, OC: $\bar{x} = 8.0$). This difference in individual prediction change between markets is also reflected by differential increases in individual prediction quality, which will be more thoroughly addressed in the following section. Noise traders who participated in the market with overconfident traders improved their predictions less after the market, which suggests that they extracted less information from the market.

Ultimately, overconfident subjects' trading behavior influences the trading behavior of noise traders. This becomes particularly evident when observing changes in noise traders' trading behavior over the course of the market periods. Compared to the treated subjects' analysis, we therefore extended the trading behavior analysis for noise traders, as shown in Figure 8.5. Instead of presenting market-based average results like for the treated subjects' trading behavior, the market periods are divided into equally-sized time intervals and results are then presented per time interval. This allows us to observe shifts in noise traders' behavior, separated by LC and OC markets.

The first row depicts the overall trading behavior of noise traders and the second row illustrates noise traders' behavior in cases where they opposed previous market signals. We observe in the first column that trading activity in terms of total stocks traded is lower in OC markets at the beginning but higher at the end of the market periods. This observation is valid for overall trading activity and opposing trades.

The second graph of row (a) illustrates that noise traders start trading very large amounts of stocks per trade (LC: $\bar{x} = 33.8$, OC: $\bar{x} = 35.0$) independent of the treatment condition, and that this amount falls and stays lower, on average, in LC markets. The same observation is reflected by average stock price movements for overall trading activity. Noise traders in OC markets moved the stock prices by larger margins per trade in the majority of market period intervals.

The last two charts in row (b) depict the differences in how strongly noise traders opposed previous market signals depending on the other subjects' treatment condition. Congruent with the hypothesis, overconfident subjects' aggressiveness seemed to translate into noise traders' behavior. At the beginning of the period, when little information

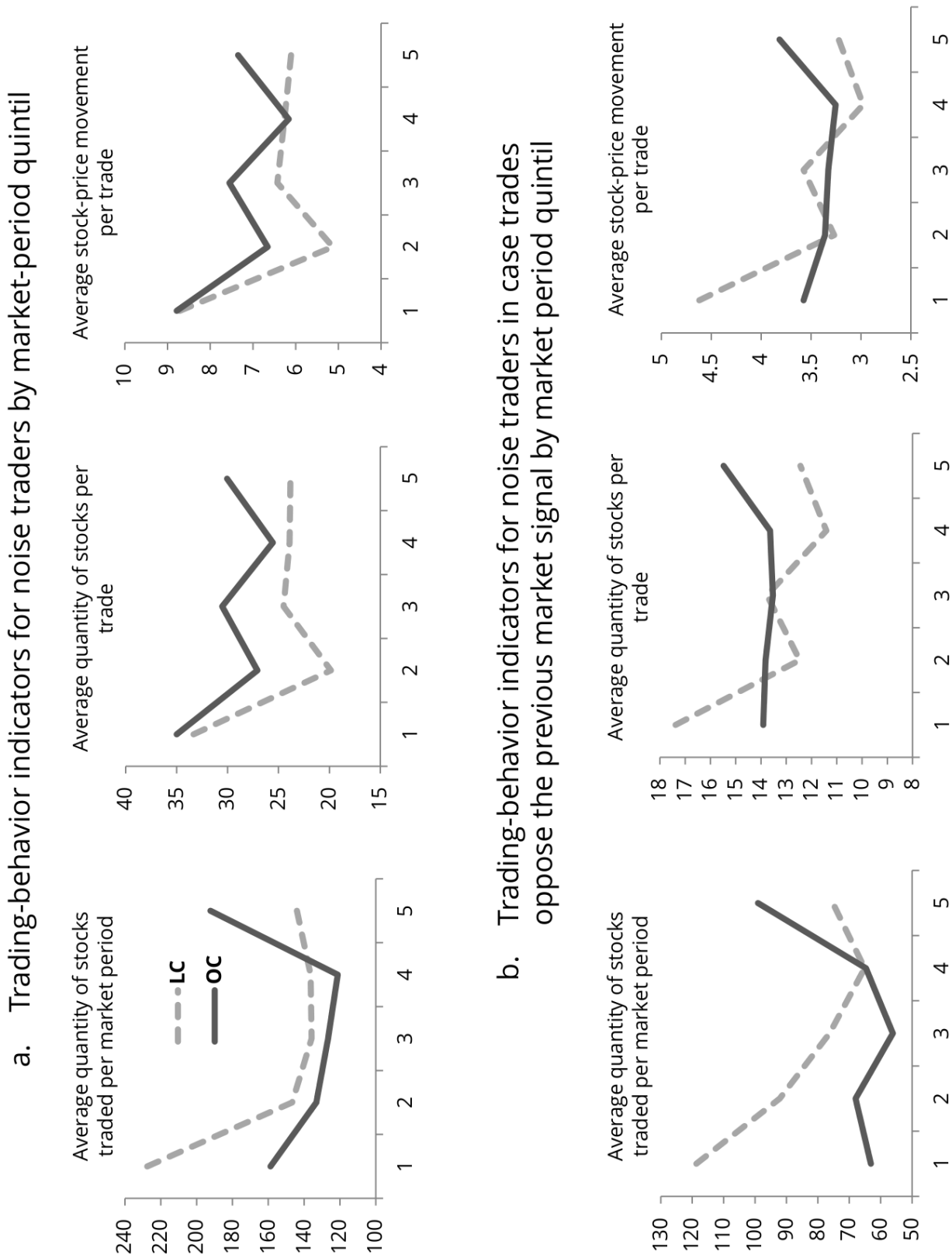


Figure 8.5.: Indicators of trading behavior by noise traders and treatment condition (Source: Own depiction)

about the trading behavior of the treated subjects was available in the market, opposing behavior was even higher in markets with LC subjects. However, as the market period progressed, the average stock quantity and price movements of noise traders' opposing trades in OC markets began to show the corresponding characteristics as in LC markets. This may indicate that noise traders learn to imitate the behavior of overconfident traders in the respective markets and to trade more strongly against opposing market signals over the course of the market period. As a consequence, all market participants would be less likely to extract information from the market; first, because market signals are less homogeneous and second, because subjects would become less willing to extract information from market signals.

Overall, the results show support for **Hypothesis 6** and the additional analysis of trading behavior indicates correspondence to the reasoning behind the hypothesis. In the given experimental setting, overconfidence negatively influenced market prediction quality because it incremented the positive relationship between noise traders' pre-market errors and market error. Overconfident subjects trade more aggressively, which prevents noise trader from learning and stimulates them to adapt similar behavior over the course of the market period.

Post-market predictions

This section presents the results regarding the effect of overconfidence on post-market prediction quality by treated subjects and the improvement in prediction quality by noise traders. We chose prediction improvement as the dependent variable for noise traders to account for the larger variance in their pre-market predictions. In contrast, manipulated subjects had lower and, importantly, more homogeneous prediction errors prior to the market. Therefore, the results can be directly regressed with their post-market prediction errors.

We first address the RSE of post-market individual predictions by treated subjects. Similar to previous analyses, we built two HLMs and included random effects variables in Model 2. Both models show that individual after-market estimates by treated subjects were significantly impacted by individual pre-market estimates ($p < 0.001$), which significantly supports **Hypothesis 7**. In addition, market errors have show no apparent significant effect on post-market prediction errors by treated individuals in the fixed- or random-coefficients models. It appears that treated subjects drew very heavily from their pre-market predictions but not from market signals to form post-market predictions. The pre-market information and subsequent individual predictions set anchors

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

Variable	Treated individuals' post-market error		Noise traders' post-market prediction improvement	
	Model 1	Model 2	Model 1	Model 2
Fixed effects				
<i>Intercept</i>	3.82*** (1.08)	3.04*** (1.07)	-1.04 (1.19)	2.18[†] (1.13)
Individual pre-market error	0.60*** (0.04)	0.59*** (0.66)	0.56*** (0,04)	0.56*** (0.06)
Market error	0.03 (0.06)	0.05 (0.04)	-0.43*** (0.06)	-0.49*** (0.07)
Random effects				
Treatment condition	-	-0.59 (0.94)	-	-0.64 (1.07)
Gender	-	0.76 (0.61)	-	-3.50** (1.46)
Financial Risk Attitude	-	0.53 (0.37)	-	0.08 (0.56)
Product Domain Involvement	-	-0.27 (0.32)	-	1.15 (1.23)
Individual pre-market error × Treatment condition	-	0.14* (0.07)	-	0.01 (0.06)
Market error × Treatment condition	-	-0.02 (0.04)	-	0.11 (0.07)
Deviance	1212.04	1186.70	1222.04	1211.94
Deviance Change (d.f.)		25.34 (11)**		10.10 (11)
Additional Variance explained		8.0%		2,9%

Notes. n for period level = 180; n for subject level = 36.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$.

Table 8.4.: The impact of treatment condition on treated subjects' post-market error and post-market prediction improvements by noise traders in the basic experimental setting

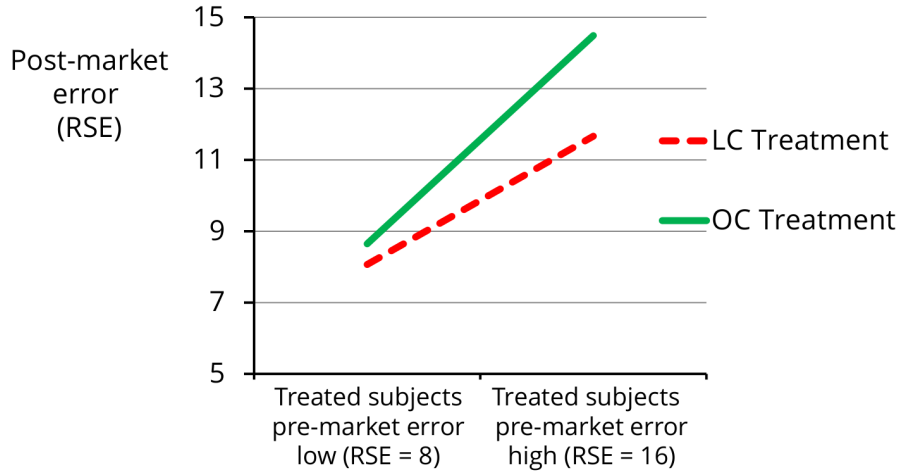


Figure 8.6.: Interaction effect of treated subjects' pre-market errors and treatment conditions on their post-market errors

with little ambiguity for the range of potentially true values that would serve as a yard stick for the validity of fluctuating market signals. Treated subjects are unlikely to process market signals to update their beliefs because market prices and their corresponding prediction errors frequently diverge from the range of potentially true values due to noise traders' activity.

Hypothesis 8 posited that overconfident individuals' post-market estimates would be more negatively affected by pre-market errors. In other words, the higher the individual pre-market errors, the worse the post-market predictions by OC participants would be. We investigated this hypothesis by looking at Model 2 random-effects results. The second column of Table 8.4 shows that the interaction effect of treatment condition and individual pre-market error is significantly positive ($p < 0.05$), which lends support to the corresponding hypothesis. This finding is visualized in Figure 8.6. By drawing from Model 2 estimators, the figure shows a stronger increase in post-market errors with increasing pre-market errors by OC subjects, as compared to LC treatment. Overconfident subjects appeared to learn less during the market to improve private predictions, which was particularly true when pre-market errors were relatively high.

An interesting question is why particularly low confidence subjects did not generate worse after-market than pre-market predictions in the case of low pre-market errors. One argument could be that the previously received information prevented them from

updating their beliefs for the worse. Erroneous market signals do not resemble predictions within the boundaries of previously internalized information.

Overall, the second model reduced variance by 8.0 percent, which translates into a significant increase in model fit ($p < 0.01$). None of the control variables in the random effects model yielded significant relationships with the individual post-market errors.

Finally, we place our focus upon the potential prediction improvements by noise traders. The results are provided in the last two columns of Table 8.4. Model 2 did not significantly improve model fit, which allows us to draw our conclusions from the standard regression Model 1.

The first model predicts that noise traders' post-market prediction improvement significantly depends on individual pre-market errors and market errors. In line with **hypothesis 9** higher pre-market errors give more room to prediction improvements by noise traders after the markets have ended. An increase of pre-market RSE by one results in an estimated increase of post-market prediction improvement of 0.56 ($p < 0.001$). We thus cannot refute **Hypothesis 9**.

Second, we observe that prediction improvement significantly decreased with increasing market error ($p < 0.001$). An increase in market RSE by one point results in an estimated decrease in post-market prediction improvement of -0.49 ($p < 0.001$). Noise traders learned from market signals and improved their predictions after the markets. The higher the market error, the less noise traders improved upon their pre-market predictions, which lends support to **Hypothesis 10**.

8.4.2. An additional setting where information for treated individuals was provided at a cost

This section focuses on the results of the additional experimental setting. This setting aimed to provide another facet to increase the reliability of the results in the face of a more realistic scenario. Instead of providing diagnostic information free of charge, treated subjects were now allowed to purchase up to three pieces of diagnostic information for a deduction in potential prize money. Accordingly, this setting also addressed the impact of treatment condition on information acquisition behavior and how potential variance in available information impacts individual predictions by treated subjects, market predictions, and post-market predictions by noise traders. We begin this analysis in the following subsection by investigating the impact of treatment condition on pre-market behavior.

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

Variable	Individual information		Treated individuals' pre-market error	
	Model 1	Model 2	Model 1	Model 2
Fixed effects				
<i>Intercept</i>	1.66*** (0.30)	0.16 (0.35)	18.34*** (1.29)	19.63*** (1.72)
Individual information	-	-	-2.77*** (0.72)	-2.94*** (0.74)
Random effects				
Treatment condition	-	1.31** (0.43)	-	-1.94 (1.61)
Gender	-	0.86† (0.44)	-	0.21 (1.26)
Financial Risk Attitude	-	-0.08 (0.40)	-	1.27 (0.94)
Product Domain Involvement	-	0.03 (0.37)	-	-0.36 (0.86)
Individual Information × Treatment condition	-	-	-	0.72 (0.73)
Deviance	137.67	117.69	1232.04	1228.81
Deviance Change (d.f.)		19.98 (4)***		3.24 (5)
Additional Variance explained		24.2%		1.4%

Notes. n for period level = 160; n for subject level = 32.

***p < 0.001; **p < 0.01; *p < 0.05; †p < 0.1.

Table 8.5.: Regression results regarding individual post-market prediction errors by treated subjects and individual prediction improvements by noise traders

Pre-market predictions

Subjects were first required to decide how much diagnostic information they wanted to pay for to potentially improve individual predictions. In the first two results columns of Table 8.5, we show how overconfidence treatment impacted individual information-buying.

As the experiment only presented three options for buying information, the HLM analysis is carried out via an ordinal regression. Ordinal regression results in probability estimates for subjects choosing one of the options. The intercept captures the overall log-odds if all predictors are controlled for at the grand-mean level. The results can be interpreted by transforming the resulting log-odds to probability $P(i)$ that an option i is chosen,

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

e.g. as in Gracia and Herrero (2008). The probability for the baseline choice (zero pieces of information purchased) are given in the first formula below. The second formula returns the overall probability for buying one piece of information by adding a threshold estimate δ . The probability of purchasing all possible pieces of information results from the preceding estimates:

$$P(0) = \exp(\text{intercept}) / (1 + \exp(\text{intercept}))$$

$$P(1) = \exp(\text{intercept} + \delta) / (1 + \exp(\text{intercept} + \delta))$$

$$P(3) = 1 - P(1) - P(0)$$

The regression produced a δ estimate of 1.72. Using the results provided in Table 8.5 and the threshold estimate, we find that the probabilities were $P = 54\%$ for purchasing zero information pieces, $P = 13\%$ for purchasing one piece of information, and $P = 33\%$ for purchasing three pieces of information when controlling for all other potential predictors.

Similar to the intercept, the other predictors can be included to calculate probability changes or they can be interpreted directly from the results table. After including the random effect variables in Model 2, we find a significant impact of treatment condition on the likelihoods for each information acquisition choice ($p < 0.01$). The positive sign indicates that overconfidence treatment increases the probability that subjects will choose the first category (i.e. buy no additional information). The model predicts that *ceteris paribus*, 81% of subjects with overconfidence treatment will not purchase any diagnostic information. At the same time, overconfidence treatment decreases the probability of subjects buying one or three pieces of information.

After receiving purchased information, the treated subjects gave individual estimates prior to the market starting. In the next step, we investigate how the amount of acquired information impacted individual pre-market prediction error. The last two columns of Table 8.5 show the regression of pre-market errors by treated individuals on the amount of information acquired, treatment condition and control variables. The deviance change from Model 1 to Model 2 shows that the random effect variables do not contribute to explaining individual pre-market errors. Only the intercept constant ($p < 0.001$) and the amount of acquired information ($p < 0.001$) significantly impacted the dependent variable. The more information an individual acquires, the lower his individual prediction error will be. Model 1 estimates that each piece of diagnostic information lowers

Variable	Market error	
	Model 1	Model 2
Fixed effects		
<i>Intercept</i>	4.45 (4.40)	5.44 (3.03)
Market information	-0.80 (0.68)	- 0.71 * (0.32)
Group Error Treatment	0.17 (0.24)	0.07 (0.21)
Group Error Noise	0.47 ** (0.17)	0.15 * (0.15)
Random effects		
Treatment condition	-	-0.85 (3.03)
Market information × Treatment condition	-	- 0.97 ** (0.32)
Group Error Treatment × Treatment condition	-	-0.02 (0.21)
Group Error Noise × Treatment condition	-	0.28 † (0.15)
Deviance	601.89	575.76
Deviance Change (d.f.)		21.60 (14)*
Additional Variance explained		9.9%

Notes. n for period level = 80; n for market level = 16.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.1$.

Table 8.6.: Impact of the presence of overconfident individuals on market prediction error when information comes at a cost

individual prediction error by 2.77 points.

Market predictions

Compared to the basic experimental setting, the overall amount of acquired information is now introduced as an independent predictor variable to the regression models that investigate the influence of overconfident market participants on prediction market error. Table 8.6 shows the results of the HLM analysis related to the market error. Observing Model 1 first, we find that average noise group errors significantly increased prediction errors ($p < 0.01$). This positive relationship stays intact when introducing random effect predictors in Model 2, yet is less significant ($p < 0.05$). Overall, the second

model explains 9.9% more variance than Model 1, which significantly improves model fit ($p < 0.05$). We now find that available market information significantly reduced prediction errors ($p < 0.05$). Each additional piece of diagnostic information reduced prediction market error by 0.71 points when controlling for the other predictor variables. Model 2 additionally reveals an interaction effect according to which an increase in market information will be of greater benefit in the case of overconfidence treatment ($p < 0.05$). Each additional piece of diagnostic information reduced prediction error in markets with overconfident subjects 0.97 points more than in markets with low-confidence subjects. This finding is interesting because the previous experimental setting did not reveal a stronger positive effect on market error by well-informed, overconfident subjects. On the contrary, market error in the presence of well-informed treated subjects was higher in OC conditions than in LC conditions.

However, this interactive relationship appears to be offset by the significantly lower total amount of information in markets with overconfident subjects. On average, the overconfident markets yielded a prediction error of 17.50 points ($sd = 12, 60$) and the low-confidence markets yielded 13.08 points ($sd = 13, 42$) prediction error.

Post-market predictions

The last section addresses individual post-market estimation errors in the setting where information comes at a cost.

Table 8.7 shows the individual post-market errors by treated individuals. Similar to the basic experimental setting, individual prediction errors significantly contributed to post-market errors ($p < 0.001$). The table additionally reveals that market errors significantly and positively impacted post-market individual errors by treated individuals ($p < 0.001$). Based on the effect-size estimations in Model 2, each point increase in market error increased individual post-market errors by 0.37 ($sd = 0.07$) points. Market error significantly influenced treated subjects' predictions after the market, while these were not significantly influenced by market error when the subjects were well-informed prior to the market starting.

Furthermore, the control variables of Gender and Product Domain Involvement had a positive influence on post-market errors. Male subjects had higher post-market errors than female subjects on average, and higher product domain involvement led to higher post-market errors. While it may seem counterintuitive at first sight, higher domain involvement could positively influence post-market errors because it might relate to less-willingness to learn from others' signals, in addition to the treatment condition.

Variable	Treated individuals' post-market error		Noise traders' post-market prediction improvement	
	Model 1	Model 2	Model 1	Model 2
Fixed effects				
<i>Intercept</i>	−0.14 (1.22)	0.91 (1.26)	−1.04 (0.87)	−1.35 (1.11)
Individual pre-market error	0.59*** (0.08)	0.52*** (0.08)	0.50*** (0.07)	0.49*** (0.05)
Individual information	0.09 (0.43)	−0.84 (0.57)	-	-
Market error	0.35*** (0.08)	0.37*** (0.07)	−0.35*** (0.06)	−0.32*** (0.05)
Random effects				
Treatment condition	-	0.31 (1.13)	-	1.32 (0.80)
Gender	-	1.42* (0.54)	-	0.16 (0.89)
Financial Risk Attitude	-	0.09 (0.59)	-	−0.42 (0.49)
Product Domain Involvement	-	2.00** (0.51)	-	−0.84 (0.69)
Individual pre-market error × Treatment condition	-	−0.08 (0.08)	-	−0.14* (0.05)
Individual Information × Treatment condition	-	−0.43 (0.42)	-	-
Market error × Treatment condition	-	0.02 (0.07)	-	0.03 (0.05)
Deviance	1121.08	1073.53	1097.72	1083.18
Deviance Change (d.f.)		47.56 (16)***		23.65 (5)***
Additional Variance explained		39.5%		7.8%

Notes. n for period level = 160; n for subject level = 36.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.1$.

Table 8.7.: The impact of treatment condition on post-market error by treated subjects and post-market prediction improvements by noise traders when information comes at a cost

We ultimately address the improvement between pre- and post-market estimates by noise traders. First, we find from observing the last two columns of Table 8.7 that intercepts are not significant either model. However, similar to the basic experimental setting, higher pre-market errors and lower market error are positively related to improvement. The more poorly noise traders estimated before the market, the more they improved their estimations after the market ($p < 0.001$), and the worse market the error, the less likely the noise traders were to improve estimations after the market ($p < 0.001$).

We finally observe an interaction effect of pre-market errors by noise traders and treatment condition. This interaction shows that noise traders improve less with increasing pre-market errors if they participate in a market with overconfident participants. Model 2 in the last column of Table 8.7 predicts that each point increase in pre-market error by noise traders leads to 0.28 points $((1) \times (-0.14) - (-1) \times (-0.14))$ less reduction in prediction error if the noise traders interacted with overconfident subjects ($p < 0.05$). This is interesting because a similar effect was not observed in the basic experimental setting. This may point to a potentially negative effect between lacking information and overconfident traders' behavior. Especially when non-treated subjects have high pre-market errors, the presence of overconfident subjects negatively influences their prediction improvements, even when controlling for higher market error.

8.5. Discussion

The second experiment focused on the impact of overconfidence on the prediction quality of information markets and its participants. We contrasted a condition in which overconfident subjects participated in information markets with a condition in which subjects with low confidence entered similar information markets. We tested two different experimental settings regarding the information base of these subjects. In the basic experimental setting, treated subjects received diagnostic information free of charge. This allowed us to study the impact of overconfidence on market prediction quality while controlling overconfident' subjects information levels externally. We then extended the basic experimental setting with a scenario in which treated subjects had the opportunity to purchase diagnostic information for a deduction from potential winnings. In all conditions and experimental settings, two treated subjects were accompanied by an equal number of noise traders, neither of whom received treatment or diagnostic information. The overall results for both scenarios will be presented and discussed in the following paragraphs.

Treatment Condition	Low confidence		Overconfidence	
Subject group	Treated	Noise	Treated	Noise
Average individual	12.42	18.46	12.82	18.64
pre-market RSE	(11.86)	(12.64)	(11.57)	(13.64)
Information market	9.21		16.37	
RSE	(6.47)		(10.29)	
Average individual	10.90	13.65	12.81	15.91
post-market RSE	(9.57)	(10.02)	(11.00)	(10.31)

Notes. n for individual RSE = 180; n for information market RSE = 90.

Table 8.8.: Individual and market prediction errors in the basic experimental setting

We summarize the prediction results for the basic experimental setting in Table 8.8. This table supports the findings of previous HLM analysis that confidence is unrelated to the quality of individual predictions when subjects are supplied with diagnostic information free of charge at the moment they receive the prediction task. As they are informed that the information is diagnostic for the underlying prediction task, subjects in both treatment conditions apparently utilized it to form individual estimations and yielded similarly low pre-market errors (12.42 for LC and 12.82 for OC).

However, we find that the prediction quality of information markets nonetheless suffered from the presence of overconfident subjects, even though their pre-market estimations did not significantly differ in quality from low-confidence subjects. As presented in the results sections, our findings support recent experimental results. Subjects who learn from the signals of overconfident individuals may themselves produce signals with higher prediction errors (Radzevick and Moore 2011). The detrimental impact of noise traders' pre-market estimations appears contingent upon the presence of overconfident traders. The statistical analysis of the experimental results reveals that increasing individual prediction errors by noise traders have a stronger negative impact on the prediction quality of market prices if overconfident traders are present. Our experimental results provide two potential explanations that could independently explain these results.

First, noise traders may learn less from overconfident traders' market signals because they align expectations less during trading. Overconfident traders oppose market signals more strongly than low-confidence treated subjects. Our results align with Odean's (1998) model, which shows that overconfident traders increase market volatility, and we argued that traders consequently update private estimations less based on market signals. These findings receive further support when observing post-market predictions by treated individuals. Overconfident individuals appear to learn (or want to learn)

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

relatively less from the information market, since individual pre-market errors attribute significantly more to their post-market errors compared to subjects in the low-confidence treatment.

Second, noise traders appear to adapt to overconfident subjects' *aggressive style of trading*. During the course of trading, noise traders oppose market signals more strongly in OC markets than in LC markets. Though the post-market predictions show no significant direct impact of treatment condition on the prediction improvements by noise traders, the experimental setup prevents us from knowing whether the effect is indirectly mitigated via the positive influence of lower market error in low-confidence markets.

We used descriptive analysis to support the hypothesis that overconfidence will have a negative influence on market prediction error. The question, however, calls for further analysis to more closely study the direct effects of overconfident subjects' actions on noise traders' reactions. In our experiment, such reactions could not be observed individually, as they were interwoven with other market actions and reactions. Future experiments can greatly support the given findings by providing a more focused view of the isolated relationship between overconfident subjects' actions and the reaction of overconfident or well-calibrated subjects.

Finally, we observe that the market mechanism achieves the transfer of diagnostic information from treated participants to noise traders. We find a significant improvement among noise traders' post-market estimations compared to their pre-market estimations, which is positively influenced by low market error and high pre-market error. In contrast to treated subjects, for noise traders, a significant share of market signals contributes to their formation of post-market beliefs. Noise traders form better private predictions after the market than before, and this effect is positively influenced by market prediction quality.

In the additional experimental setting, subjects decided whether they wanted to acquire information at a cost. We summarize the average prediction errors by treatment condition in Table 8.9. The first row of results shows the average amount of diagnostic information acquired by treated individuals, which preceded the pre-market estimations. Here, our results support previously discussed findings that overconfident subjects will engage in less information acquisition because the information appears less valuable to them, compared to subjects in the low-confidence condition. This translates into higher individual prediction error in subjects who received overconfidence treatment. Accordingly, market prediction error is, on average, higher in markets with overconfident subjects than in those with subjects from the low-confidence condition.

8. Experiment 2: Overconfidence and the Prediction Quality of Information Markets

Treatment Condition Subject group	Low confidence		Overconfidence	
	Treated	Noise	Treated	Noise
Average amount of information bought	1.67	–	0.55	–
Average individual pre-market RSE	15.06 (13.11)	18.34 (12.16)	16.05 (10.48)	18.97 (12.61)
Information market RSE	11.60 (11.77)		16.87 (12.22)	
Average individual post-market RSE	13.91 (13.50)	13.76 (9.87)	14.57 (10.48)	17.36 (12.30)

Notes. n for individual RSE = 160; n for information market RSE = 80.

Table 8.9.: Individual and market prediction error in the additional experimental setting

However, an interesting finding is that additional piece of diagnostic information reduces prediction error in markets with overconfident subjects more so than in markets with low-confidence subjects. This finding offsets the results from the first experimental setting, in which markets with overconfident subjects yielded higher predictions error, even though treated subjects always had access to three pieces of diagnostic information.

One potential explanation could be that OC subjects who purchased the maximum amount information were least affected by the treatment. Overconfident subjects who purchased extra information did not acquire information because their environment signaled them to do so, but because the need emerged from within. Hence, their trading behavior may correspond much less to that of the overconfident trading behavior that explained high market errors in the previous experimental settings. As a consequence, subjects who feel the urge to purchase information, even after being informed about their superior prediction quality, may be self-selecting into markets in which participants are more open to valuable external information. Thus, markets with overconfident subjects may benefit more strongly from additional information than in similar markets with low-confidence subjects.

Ultimately, the market prediction errors are consistent with the findings of the HLM analysis that overconfident subjects' reluctance to acquire information apparently outweighs the fact that the presence of overconfident subjects induces lower market error at comparable (and sufficiently low) information levels. There are too few instances in which overconfident subjects enter markets similarly well-informed as subjects in the low-confidence condition, to potentially compensate for the effect of lacking information. In summary, both experimental settings showed that the presence of overconfident par-

8. *Experiment 2: Overconfidence and the Prediction Quality of Information Markets*

ticipants can have a negative impact on the quality of information market predictions. When other subjects possess relatively little information about the prediction task, the presence of overconfident subjects will decrease the likelihood that these subjects will absorb information from overconfident subjects via the market's trading mechanism. This leads to higher market error and higher individual prediction errors for uninformed traders.

This may be particularly important to consider in the light of heterogeneously dispersed information domains among market participants, which could be common in the context of innovation evaluation. In cases where overconfident subjects are present and possess diagnostic yet independent information from different domains, overconfident trading-behavior may prevent the effective aggregation of this information. When overconfident participants possess specific expertise, the remaining participants become noise traders who obstruct effective integration of the overconfident subjects' expertise. This will negatively influence the quality of market predictions and individual post-market predictions by participants and other groups or individuals who draw diagnostic information from these predictions, such as innovation management and planning staff.

Part IV.

Synthesis

9. Synthesis

9.1. Summary of results

The main goal of this thesis was to create a more thorough understanding about the impact of judgmental biases on the evaluation of innovation via information markets. This goal was split in two research objectives.

In Part Two of this thesis, we addressed the **first research objective** and created a conceptual background that set the scene for subsequent empirical analysis. A thorough introduction to information markets for innovation evaluation and judgmental biases prepared a common and detailed understanding of the research context. We found that judgmental biases are frequently present in innovation management decision-making and may significantly impact the outcomes of innovation evaluation tasks. Of the most prevalent biases identified in Section 4.3, overconfidence appears to be particularly important in the context of innovation evaluation in general and information markets in particular.

From there, we determined our **second research objective**, which was to empirically study the impact of overconfidence on the evaluation of innovation via information markets. First, the empirical focus was narrowed to the potential impact of overconfidence on the quality of the innovation evaluation. We defined overconfidence as the individual tendency to overestimate own prediction performance overall and in relation to others. Two experiments with information markets were conducted.

The **first experiment** investigated the impact of overconfidence on individual behavior in the context of information markets. We drew inferences about individual behavior by creating an information market environment that allowed us to analyze individual behavior in a controlled setting. Subjects received a treatment that manipulated their confidence levels, after which they traded with a controlled artificial market agent over a number of sessions. The first experiment demonstrated that higher levels of confidence

9. Synthesis

induce individuals to trade more stocks in information markets, engage in trading earlier, trade more in opposition to other traders' deflecting beliefs, and be less likely to update personal beliefs in relation to market signals after the information market has finished. These results indicate a stronger influence of overconfident subjects' beliefs on the aggregated information that is expressed via market prices.

The **second experiment** investigated the impact of overconfidence on the prediction quality of information markets and their participants. Subjects received a treatment that manipulated their confidence to either overconfident or low confidence levels. Pairs of similarly treated subjects were brought together with two noise traders who had not received any treatment. Thus, the second experiment featured information markets in which four traders aggregated their expectations via trading. The results documented that the presence of overconfident participants may significantly lower the prediction quality of information markets. When overconfident subjects are **well-informed** regarding the underlying prediction target, they exhibit strong convictions regarding their interpretations of the provided information. Consequently, they trade more aggressively, as suggested by the first experiment, and more strongly oppose other participants' market signals. This limits both participants' learning from those market signals and efficient information aggregation. In the presence of overconfident traders, market predictions are less likely to converge to form a homogeneous signal that could be better interpreted by non-informed subjects to allow them to update private information. In markets with overconfident subjects, noise traders' predictions have a more strongly negative impact on market prediction quality and these traders will be less likely to improve their private predictions after the information markets have finished.

When overconfident subjects can **acquire diagnostic information** prior to the market, they will be less likely to do so, since acquiring information is less valuable to them. Acquiring less information leads to higher individual prediction error on the part of overconfident subjects before the information markets and less information in the information market, both of which ultimately negatively impact market prediction quality. Higher prediction errors can be related to (1) less trading activity by overconfident subjects because they are aware that they are less likely to acquire information and thus less informed and (2) less informed trading activity by overconfident subjects if they ignore the fact that they are less informed because they see less value in acquiring information.

The two experiments expand the current state of research with regard to judgmental

biases in the context of information markets. The results of the first experiment provide a detailed picture of the impact of overconfidence on individual behavior in a controlled laboratory experiment. The second experiment demonstrates that the presence of overconfident subjects will systematically lead to lower prediction quality in information markets in the given informational environments.

9.2. Limitations

Behavioral experiments are narrowly focused with regard to incentives, participants, inter-subject relationships, and experimental environments to allow the investigation of causal relationships while controlling for potential interference. Such rigorous control over research conditions produces numerous limitations, which potentially reduce the extent to which the findings can be applied to relationships in related experimental or real-world contexts. Thus, it is important to investigate the limitations of the present experimental research for two reasons. First, awareness about the limitations of an experiment provides the necessary consideration for developing theoretical implications from the results. Researchers are stressed not to overly generalize from their findings. Second and closely related to the first point, theoretical assumptions to which empirical results may appear ambiguous provoke suggestions for future research to increase the findings' reliability and obtain a clearer understanding about the underlying mechanisms that produce the phenomena in question.

First, all of our experimental **subjects were drawn from a body of master students** in engineering disciplines. Our sampling therefore partly confounds our inference from the current body of research that overconfidence is most likely to be present among entrepreneurs, innovation managers, and executives. While entrepreneurs and managers in innovative technology companies may be more likely to possess engineering degrees than the general population, certainly not all the students in our sample will join these ranks. However, we addressed this issue by developing an experimental treatment that allowed us to create overconfidence artificially. On the one hand, such an experimental treatment comes with a significant benefit. Overconfident subjects can be allocated in any fashion that the experiments require. On the other hand, such treatment may lack external validity if the artificially-induced overconfidence differs from the natural overconfidence among the above-mentioned groups. While our experiments showed that the behavior by subjects with induced overconfidence resembled the behavior of untreated

student subjects with natural overconfidence, we cannot draw inferences about a similar resemblance to the behavior of overconfident entrepreneurs, inventors, innovation managers, or senior leadership.

Second, there may be **differences between overconfidence as induced by our treatment and natural overconfidence**. However, we did not find significant differences in the dependent variable means when comparing naturally overconfident individuals have a strong better-than-average perception in our control group and participants in the treatment group. Second, we had to rely on students as subjects for our experiment and could not recruit real managers, engineers or blue-collar workers from automobile companies. However, we have put forward arguments for why we still believe our manipulated subjects felt sufficiently involved in the industry to make informed statements. We believe that the effects we found would be even stronger among real decision-makers, as they would be able to tap a larger pool of private information when evaluating innovations.

Third, we aimed at creating relevant innovation evaluation tasks by using real-world evaluation examples. However, our innovation tasks dealt mainly with **innovations that were already entering the market**. Therefore, we can only assume that similar patterns would arise when evaluating at the fuzzy front end of innovations, e.g. during the idea evaluation process. Yet, drawing from current research, we believe that overconfidence may have a far greater impact on individual behavior at the beginning of innovation endeavors, when context is less understood and uncertainty considerably higher (Fowler and Johnson 2011). Alas, it would be very difficult to validate the evaluation outcome of ideas, since a true underlying value would not be available any time soon.

Fourth, the incentives in our experiments were fixed at amounts that - in the case of winning one of the vouchers - meant significant financial gains to our participants in relation to the time the experiments lasted. The average student salary at universities in the German state of Hamburg was €8.49 per hour at the time of the experiments and potential winnings amounted to €20 or €50. Therefore, our experiments should have provided relevant extrinsic incentives for truthful revelation of beliefs. Yet of course, **real-world information markets may differ strongly with regard to incentives**. Frequently, the relationship between potential monetary awards and regular participants' salaries is much higher in corporate information markets for evaluation of innovations. For

example, the largest German telecommunications company planned to provide a pink Volkswagen convertible as a prize to spur female participation¹. However, researchers have frequently stressed that motivation-based incentives (for having better predictions than others and achieving the higher portfolio value) may be a sufficiently strong driver for effective participation (Dahan et al. 2010).

Additionally, our results only allow limited inferences about how overconfidence may impact information markets that require subjects to **invest private funds**. While the additional experimental setting of the second experiment may provide some indication of how overconfidence interacts with the risk of losing funds, the subjects only risked “winning less” rather than leaving the experiment with fewer assets than they had possessed before starting.

Fifth, our **experimental environment** requires critical examination regarding to what extent it resembles real information markets for innovation evaluation. Any differences may need closer examination to increase reliability when making inferences about real-world information markets. To begin with, our innovation evaluation tasks were all concerning the future market success of innovations that had recently been introduced to the market. We have presented arguments for why we choose such tasks and under which circumstances we may expect similar results if the tasks concerned early stage innovation ideas or concepts. Yet our experiments cannot provide empirical support to these arguments. Closely related to the tasks, the information cues provided during the second experiment were homogeneous among participants (if they acquired equal amounts of information). Our experiment, therefore, did not investigate how the presence of overconfident subjects with very heterogeneous knowledge (e.g., marketing and R&D managers) impact prediction results. While the results of the second experiment indicate the lack of consensus building and its negative consequences on noise traders’ prediction formation, we did not vary the heterogeneity of information systematically to provide further insight. Also, subjects were paid out in our experiment only after the real values materialized. As our information markets dealt with predictions of future market shares, there was a significant time gap between participation and payout. Hence, our experiments did not answer the question how success (or the lack thereof) by overconfident individuals over the course of multiple rounds of payout might impact their behavior and influence in information markets. The results of our experiments should therefore be specifically regarded as discussing the impact of overconfidence on the re-

¹The authors learned this from a personal talk with the company’s project leader.

sults of first-time applications of information markets for the evaluation of innovation. Finally, our experimental information-markets ran on much smaller scale than real-world information markets. While our markets featured four participants, real-life information markets usually contain at least dozens and up to thousands of active traders, as was the case in the Google and the Hileman information markets (Cowgill et al. 2008; Othman and Sandholm 2010). Thus, our results may overstate the effect of overconfident subjects in the cases where they are relatively less present in large-scale real-life information markets. However, we presented arguments in Section 5.1 asserting that innovation evaluation tasks are likely to attract a significant proportion of overconfident subjects. Empirical insights about the true degrees of overconfidence among innovation evaluators could provide a meaningful benchmark for this experiment's results and how well the results can be projected onto real-life information markets.

Finally, subjects only predicted the success of one innovation per market period. Real-world information markets usually feature a much larger number of parallel prediction tasks, which may lead to increasing presence of overconfident subjects as they self-select to trade stocks for those tasks, whose future they feel they know most about (Camerer and Lovo 1999). There is also some evidence that subjects may adjust their confidence after receiving feedback on their real performance (Mahajan 1992). Still, field data from financial investors and entrepreneurs indicates that overconfidence is likely to persist after subjects have received valid feedback (Glaser et al. 2003; Oberlechner and Osler 2012).

9.3. Implications for innovation management

The given research provides implications for practitioners in innovation management. In this section, we draw from the results of the conceptual and the empirical portions of this thesis to highlight meaningful learning for management practice.

The **conceptual part** of this thesis first focused the relevance and consequence of uncertainty in the context of innovation. Innovation managers need to be aware that innovation development and innovation evaluation are characterized by inherent uncertainty. Innovations do not allow the innovator to determine market success *ex ante* without error. Factors internal and external to the innovating organization influence the uncertainty of the innovation. Externally, uncertainty will be higher in situations in which

market-related information is hard to acquire due to lack of availability or is dynamically changing and therefore less diagnostic about relevant future market states. Internally, uncertainty may be positively related to a lack of inter-departmental communication, which weakens the interface between market-related needs and the organizational resources and capabilities to match those needs.

Innovators need to be aware that **valid innovation evaluation does not eliminate uncertainty** but can greatly contribute to reducing decision-making error in innovation development. This is particularly important during the early phases of the innovation process because the earlier errors are committed, the more likely they will be to lead to subsequent error and financial losses. We have argued that uncertainty can be best reduced by acquiring and aggregating information to validly assess an innovation's success potential. Particularly, **the recombination of heterogeneous expectations** of human informants such as customers, employees, suppliers, competitors and experts **can provide meaningful information**. Such a combination can be best carried out via interactive group methods that allow subjects to exchange information and learn from others' expectations so as to update private information. A well-designed, interactive group-method allows the integration of large numbers of relevant subjects, provides incentives that are compatible with valid innovation evaluation, and feature an effective algorithm for aggregating subjects' expectations.

We thus introduced **information markets as a well-suited method for valid innovation evaluation**. Empirical studies over the last decade have found that information markets in particular may be the best fit for the abovementioned prerequisites. These markets run on the principles of stock markets and use them to evaluate the potential outcome dimensions of innovation endeavors such as future market shares, release dates, development costs, or market prices. Outcome dimensions are rendered as derivatives so that participants can trade with regard to their expected outcome. The derivatives' market prices represent aggregated expectations because they incorporate the trading intentions of all market participants. The market prices are theoretically considered efficient aggregates of heterogeneously dispersed information (Arrow et al. 2008). Still, innovation managers must be aware that applying information markets for innovation evaluation requires careful consideration of various design options. As future innovation events may never occur, information market stocks may feature alternative payout schemes that are based on market-inherent events, such as average stock prices or last-

order prices. Resulting payouts do not necessarily need to be related to financial rewards. We have documented that even non-monetary rewards can produce sufficiently good results.

While **human informants** appear to be very valuable information sources for evaluation of innovation, their contributions **will often be impaired by judgmental biases**. Managers of group-based innovation evaluation processes need to be aware that human informants may access, process and distribute information in a way that is unrelated to the fundamental goal of valid innovation evaluation, even if the informants might not be aware of such action. Such action occurs because informants apply heuristics that bias perceptions, signal use, or exhibit motivations that are not aligned with evaluation goals. Our literature review highlighted that false perceptions about the representativeness of local signals, confirmation seeking, loss aversion, status-quo preference and overconfidence are particularly common biases in the context of innovation management. **Overconfidence, in particular, can harm valid innovation evaluation.** Overconfidence is highly prominent among subjects in the context of innovation evaluation because it is closely related to the risk seeking that enables subjects to embrace innovation in the first place. Overconfidence has also been related to overly optimistic resource allocation in innovation endeavors. Related domains such as financial market research or entrepreneurship show that overconfident subjects will often undermine the validity of using public signals over private information. This negatively influences the outcomes of group-based innovation evaluation because the overconfident individuals do not sufficiently absorb valid information.

The empirical part of this thesis provides a more detailed analysis of how the presence of overconfident subjects may influence the results of information markets. Even though information markets can be regarded as a relatively well-suited instrument to validly assess the success potential of innovations, researchers have previously stressed the importance of evaluating their robustness against participants' overconfidence (Wu et al. 2008).

The first experiment showed that the level of individual confidence may significantly impact trading behavior in information markets. For innovation management, this means that observations from related domains regarding the behavior of overconfident subjects in market-based environments will translate into behavior within information markets.

9. Synthesis

According to the empirical results, overconfident subjects will be more likely to be the first to articulate opinions on evaluation outcomes. As a consequence, both the content of their signals and the style of their signaling may set primary anchors for how the other subjects engage in the group-based evaluation mechanism. Furthermore, overconfident subjects show more aggressive trading behavior, which manifests in more trading overall, more trading based on initial beliefs, and less willingness to update beliefs according to aggregate market information after the markets have ended. Overconfident subjects are more likely to oppose market signals that oppose their initial beliefs. They are less likely to extract information from these market signals to update private expectations, which influences both trading behavior and post-market beliefs. This implies that overconfident subjects learn less from others in group-based innovation evaluation via information-markets. As already discussed, decision-makers, who are likely to be overconfident in the innovation context, may neglect to aggregate group expectations, even if they have participated in the aggregation mechanism. This could ultimately counteract the main reason why interactive group-methods are applied in the first place. Such implications are important to highlight along with the consequences of overconfident subjects' behavior, which were analyzed in detail in the second experiment. Overconfident subjects would be hard to observe *in vivo* in an information market. Market actions cannot usually be attributed to the actions of certain subjects via observing market signals, as opposed to the other aggregation mechanisms discussed, such as FTF meetings or Delphi. In the case of information markets, anonymity hinders insight into the impact of subjects' biases on the validity of their actions.

Focusing the outcomes of innovation evaluation via information markets, the second experiment demonstrates that the presence of **overconfident subjects** and their subsequent behaviors **have a harmful impact on the prediction quality the markets**. In particular, the presence of overconfident subjects has a significantly negative effect on the prediction quality of information markets in the context of innovation evaluation. When participants are well-informed regarding the underlying prediction targets, overconfident subjects make equally good private predictions but their particular behavior prevents effective aggregation of individual beliefs on the market. Their aggressiveness appears to transmit to other subjects and subsequently prevents those subjects from learning as well from market signals as if overconfident subjects were not present. In the case where information levels are not high *per se* but can be privately increased via information acquisition, overconfident subjects are less likely to acquire information

and will thus enter the market less well-informed. While this may reduce their aggressiveness, the lack of information again has a negative influence on aggregate prediction quality of the market.

Do our results imply that innovation managers should steer clear of applying information markets to evaluate the success potential of innovations? Our analysis only looks at the results of information markets with overconfident subjects as compared to without overconfident subjects. Thus, our experiments do not allow inferences about how overconfident subjects may impact the results of competitive methods for information aggregation, such as the face-to-face meetings or Delphi methods. We have argued that overconfident subjects are very likely to be always present in the context of innovation evaluation. Therefore, previous field studies, that have compared the performance of information markets to similar methods, have likely been subject to the influence of overconfidence. The competitive advantage of information markets over alternative methods would hence indicate that overconfidence equally (or more strongly) influences these competing methods.

However, the results do imply that initiators of information markets should be aware of the potential threat to evaluation quality the presence of overconfident subjects poses. Most importantly, our findings allow us to suggest three measures with which to act upon such awareness.

First, initiators of information markets for innovation evaluation would be well advised to **confront (overconfident) subjects before the market regarding potential biases** to increase prediction quality. In our confidence treatment, the negative feedback condition was more valid compared to the positive feedback when using average and true underlying performance as an indicator. In the second experiment, market predictions were significantly better in those markets in which subjects were given negative (or more valid) feedback before the market. This may have increased their openness to use external signals to form private predictions. For example, the low-confidence subjects were more willing to acquire costly information before the market's start. In addition, subjects that were confronted with potential gaps in performance compared to peers were more willing to learn from aggregate expectations via market prices. This appears to have had a positive effect on their private expectations after the market and on the behavior of other subjects who contingently exhibited more willingness to learn from market signals. As a consequence, confronting subjects with performance gaps

may increase the quality of private as well of aggregate group predictions when applying information markets to evaluate innovations.

Second, initiators should carefully **evaluate the benefits and pitfalls of complete anonymity for information-market participants**. On the one hand, it may be helpful to publicize private and average group-based estimations before the market to raise subjects' awareness about the distribution of beliefs and the context of their private predictions. For example, participants could be asked to connect each trade with a short text message about why they made that trade so that other subjects could better reflect upon market actions. On the other hand, such action could produce fraud to maximize personal profits and prevent valid belief revelation in the case that organizational or social issues prevented subjects from publicly stating their true private beliefs.

Finally, initiators would well advised to **run more than one information-market sequence, including payouts**. As many researchers in the domain have already pointed out in different contexts, individuals who are most beneficial to valid innovation evaluation are also most likely to benefit personally if incentives in information markets are aligned (Spann and Skiera 2003b; Dahan et al. 2010). Personal profits grow as long subjects contribute to more accurate market predictions. However, initiators should be careful to rely on additional sequences to counteract the negative presence of overconfident subjects. The second experiment demonstrated that other subjects' behavior can be contingent upon overconfident subjects' behavior. Hence, not only could overconfident subjects provide less valid information but influence the performance of all market participants, which would reduce the incentive effect for rooting out the harmful influence of overconfident subjects over the course of multiple market sequences.

To summarize, research has shown that innovation evaluation is crucial for increasing the success potential of innovations. Information markets appear to be a well-suited method for sourcing and aggregating valid information to increase evaluation quality. Still, practitioners in innovation management need to be aware that judgmental biases in general, and overconfidence in particular, may negatively influence the quality of information market predictions. Our current research has led to three suggestions to potentially reduce the negative influence of overconfidence. First, overconfident subjects should be exposed to their level of confidence prior to participating in information markets. Second, some information about belief distribution should be shared among information market participants. Finally, information markets should provide incentives that

align with valid information revelation. These markets can minimize the influence bias with each sequential application, as under-performing subjects are increasingly likely to remove themselves from the evaluation process.

9.4. Research outlook

Our research has focused on a very specific topic of information markets in the context of judgmental biases and innovation evaluation. Such specialization provides fruitful avenues for future research. Content and design choices allow researchers to provide valid results in answer to the research questions but also limit the scope of the research. In addition, such results can provoke novel questions that may be addressed by future research.

Our literature review provided an overview about current research on the impact of biases on decision making in innovation management, with a particular focus on innovation evaluation. While there exists extensive research in this domain, we also find many aspects that are well-suited for further research. In particular, the literature review showed that while research on judgmental biases in innovation management has provided much insight into how certain biases act, much less research has been done on fundamental questions such as why decision makers in innovation management are often subjects to biases and how these biases materialize on an individual level. We found little research exploring the psychological mechanisms that explain the origins of innovation-related biases and relate them to the outcomes of innovation endeavors. It appears that much room still exists to expand the phenomenological research on bias impact based on theoretical psychological research in the domain of bias origins. This may provide a better understanding about the mechanisms that lead to biases in innovation management and, from a practical standpoint, assist in developing valid countermeasures to increase decision quality.

Furthermore, the literature review targeted a broader field of decision making in innovation management but this study focused on a narrower area of innovation evaluation via information markets. To start with, we did not find much research that has narrowly addressed the impact of biases in the context of evaluating the success potential of innovations. Our research only addresses the impact of overconfidence with regard to a very particular method for innovation evaluation. There is much room to study the impact

9. Synthesis

of further biases on the outcome of innovation evaluation via information markets or further methods for innovation evaluation. Endowment effects, and subsequently, loss aversion, may be particularly interesting to study in the context of innovation evaluation as they are attributed as having much influence in the perseverance of inventors or entrepreneurs when objective considerations suggest divesting.

Our treatment check has shown that subjective confidence levels can be successfully manipulated in the context of innovation evaluation. Yet, the treatment check was only applied to two technological innovations that were already close to market entry. Whether such a treatment check is also feasible with less well-defined tasks requires further investigation. Such a treatment would more strongly address innovation evaluation at the very front end of the innovation process. It may also be interesting to compare the feasibility of false-feedback-based manipulation with alternative approaches that are frequently used in behavioral manipulation studies, such as vignetting to induce judgmental biases like overconfidence.

The main empirical focus of the empirical part of this study concerned the impact of overconfidence on the behavior of information market participants and on the prediction quality of information markets for innovation evaluation. The first experiment highlighted the particular characteristics of overconfident subjects' trading behavior. Overconfident subjects trade earlier, trade more stocks, and are less willing to update their beliefs compared to their less confident peers. However, the first experiment provided a very clinical trading environment in the laboratory setting. This suggests that additional support for our findings from a more real-world setting should be pursued. The artificial agent traded "against" the subjects beliefs in a majority of instances, which likely emphasized the contrast between the actions of differently treated subjects. In a more real-world application, much more noise via more heterogeneous expectations would enter the market. More research is needed to determine whether the same trading behaviors would persist in real-world information markets that feature many more traders with more heterogeneous expectations and noisy signals.

The second experiment focused on the influence of overconfident traders on the prediction quality of the market. We found that the presence of overconfident subjects has a significantly negative impact on the prediction quality of information markets. Our research supports current findings from related domains in experimental market research (Deaves et al. 2009) and financial market research (Glaser et al. 2003). However, the

experiment created a very specific environment, which calls for further investigation in other settings that more closely resemble real-world scenarios. In our basic experimental scenario, treated subjects were informed via similar pieces of diagnostic information. Yet, as argued in the second part of the thesis, information sourcing requires the integration of informants with valuable, but at the same time heterogeneous, information. Results may differ if subjects draw expectations from different knowledge bases. On the one hand, such distribution of information may increase the presence of overconfidence because subjects would be aware of the exclusivity of their information and would put positive emphasis on the value of their private information. On the other hand, common knowledge about the importance of several domains of knowledge (which are not all available to any one participant at once) may increase the subjects' sensitivity to market signals and reduce the potentially negative direct and contagious effects of overconfidence.

Part V.

Bibliography

Bibliography

- Agnew, C. E. (1985). "Bayesian consensus forecasts of macroeconomic variables." In: *International Journal of Forecasting* 4.4, pp. 363–376.
- Ahituv, N., J. Zif, and I. Machlin (1998). "Environmental scanning and information systems in relation to success in introducing new products." In: *Information & Management* 33.4, pp. 201–211.
- Albright, K. S. (2004). "Environmental scanning: radar for success." In: *Information Management Journal* 3, pp. 38–44.
- Alicke, M. D. and C. Sedikides (2009). "Self-enhancement and self-protection: What they are and what they do." In: *European Review of Social Psychology* 20.1, pp. 1–48.
- Allison, P. D. (2001). *Missing Data*. 1st ed. Thousand Oaks: Sage, pp. 1–93.
- Alves, J. et al. (2007). "Creativity and Innovation through Multidisciplinary and Multi-sectoral Cooperation." In: *Creativity and Innovation Management* 16.1, pp. 27–34.
- Amabile, T. M. (1998). "How to kill creativity." In: *Harvard Business Review* 76.5, pp. 76–87.
- Ames, D. R. and L. K. Kammrath (2004). "Mind-Reading and Metacognition: Narcissism, not Actual Competence, Predicts Self-Estimated Ability." In: *Journal of Non-verbal Behavior* 28.3, pp. 187–209.
- Amit, R. and P. J. H. Shoemaker (1993). "Strategic assets and organizational rent." In: *Strategic Management Journal* 14.August, pp. 33–46.
- Anderson, C. and G. J. Kilduff (2009). "Why do dominant personalities attain influence in face-to-face groups? The competence-signaling effects of trait dominance." In: *Journal of Personality and Social Psychology* 96.2, pp. 491–503.
- Anderson, C. et al. (2012). "A Status-Enhancement Account of Overconfidence." In: *Journal of Personality and Social Psychology*.
- Armstrong, J. S. (2001a). "Extrapolation for Time-Series and Cross-Sectional Data." In: *Principles of Forecasting*. Ed. by J. S. Armstrong. Boston: Kluwer Academic Publishers, pp. 217–243.

Bibliography

- Armstrong, J. S. (2001b). "Selecting Forecasting Methods." In: *Principles of Forecasting*. Ed. by J. S. Armstrong. Boston: Kluwer Academic Publishers, pp. 365–386.
- Armstrong, J. S. and F. Collopy (1992). "Error measures for generalizing about forecasting methods: Empirical comparisons." In: *International Journal of Forecasting* 08, pp. 69–80.
- Arrow, K. J. et al. (2008). "The Promise of Prediction Markets." In: *Science* 320.8, pp. 9–10.
- Ashraf, N., C. F. Camerer, and G. Loewenstein (2005). "Adam Smith, Behavioral Economist." In: *The Journal of Economic Perspectives* 19.3, pp. 131–145.
- Astebro, T., S. A. Jeffrey, and G. K. Adomdza (2007). "Inventor Perseverance after Being Told to Quit: The Role of Cognitive Biases." In: *Journal of Behavioral Decision Making* 20.3, pp. 253–272.
- Babrow, A. S. (2001). "Uncertainty, value, communication, and problematic integration." In: *Journal of Communication* 51.3, pp. 553–573.
- Barber, B. M. and T. Odean (2001). "Boys will be boys: Gender, overconfidence, and common stock investment." In: *The Quarterly Journal of Economics* 116.1, pp. 261–292.
- Baron, R. A. (2000). "Psychological Perspectives on Entrepreneurship: Cognitive and Social Factors in Entrepreneurs' Success." In: *Current Directions in Psychological Science* 9.1, pp. 15–18.
- Bayus, B.L., G. Erickson, and R. Jacobson (2003). "The financial rewards of new product introductions in the personal computer industry." In: *Management Science* 49.2, pp. 197–210.
- Belleflamme, P., T. Lambert, and A. Schwienbacher (2013). "Crowdfunding: Tapping the right crowd." In: *Journal of Business Venturing*.
- Belsley, D. A., E. Kuh, and R. E. Welsch (2005). *Regression diagnostics: Identifying influential data and sources of collinearity*. Vol. 571. John Wiley & Sons.
- Ben-David, I., J. R. Graham, and C. R. Harvey (2007). "Managerial overconfidence and corporate policies."
- Benos, A. V. (1998). "Aggressiveness and survival of overconfident traders." In: *Journal of Financial Markets* 1.3-4, pp. 353–383.
- Berg, J.E., F.D. Nelson, and T.A. Rietz (2008). "Prediction market accuracy in the long run." In: *International Journal of Forecasting* 24.2, pp. 285–300.

Bibliography

- Bergstein, H. and H. Estelami (2002). "A survey of emerging technologies for pricing new-to-the-world products." In: *Journal of Product & Brand Management* 11.5, pp. 303–319.
- Bessant, J. and H. Rush (1995). "Building bridges for innovation: the role of consultants in technology transfer." In: *Research Policy* 24.1, pp. 97–114.
- Bikhchandani, S., D. Hirshleifer, and I. Welch (1998). "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades." In: *The Journal of Economic Perspectives* 12.3, pp. 151–170.
- Billeter, D., A. Kalra, and G. Loewenstein (2011). "Underpredicting Learning after Initial Experience with a Product." In: *Journal of Consumer Research* 37.5, pp. 723–736.
- Biyalogorsky, E., W. Boulding, and R. Staelin (2006). "Stuck in the Past: Why Managers Persist with New Product Failures." In: *Journal of Marketing* 70.2, pp. 108–121.
- Blohm, I. et al. (2011). "Idea evaluation mechanisms for collective intelligence in open innovation communities: Do traders outperform raters." In: *Proceedings of 32nd . . .*
- Bonabeau, E. (2009). "Decisions 2.0: The Power of Collective Intelligence." In: *MIT Sloan Management Review* 50.2, pp. 45–52.
- Booz-Allen, Hamilton (1982). "New product management for the 1980s." In: *Booz-Allen & Hanilton, New York*.
- Bothos, E., D. Apostolou, and G. Mentzas (2009). "IDEM: A Prediction Market for Idea Managment." In: *WEB 2008*. Ed. by Christof Weinhardt, Stefan Luckner, and J. Stöß er. Vol. 19. 1. Springer-Verlag, pp. 1–13.
- Boudreau, K. J., N. Lacetera, and K. R. Lakhani (2011). "Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis." In: *Management Science* 57.5, pp. 843–863.
- Brashers, Dale E. (2001). "Communication and Uncertainty Management." In: *Journal of Communication* 51.3, pp. 477–497.
- Brown-Kruse, J. L. (1991). "Contestability in the Presence of an Alternate Market: An Experimental Examination." In: *The RAND Journal of Economics* 22.1, p. 136.
- Brun, E., A. S. Saetre, and M. Gjelsvik (2009). "Classification of ambiguity in new product development projects." In: *European Journal of Innovation Management* 12.1, pp. 62–85.
- Bullinger, H.-J. (1990). *F & E heute: Industrielle Forschung und Entwicklung in der Bundesrepublik Deutschland*. gfmt.

Bibliography

- Busenitz, L. W. and J. B. Barney (1997). "Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making." In: *Journal of Business Venturing* 12, pp. 9–30.
- Caballé, J. and J. Sákovics (2003). "Speculating against an overconfident market." In: *Journal of Financial Markets* 6.2, pp. 199–225.
- Camerer, C. F. and E. Fehr (2006). "When does "economic man" dominate social behavior?" In: *Science* 311.5757, pp. 47–52.
- Camerer, C. F. and G. Loewenstein (2004). "Behavioral economics: Past, present, future." In: *Advances in Behavioral economicshavioral Economics*. Ed. by Colin F Camerer, George Loewenstein, and Matthew Rabin. 1st ed. New Jersey: Princeton University Press, pp. 1–61.
- Camerer, C. F. and D. Lovallo (1999). "Overconfidence and Excess Entry : An Experimental Approach." In: *American Economic Review* 89.1, pp. 306–318.
- Carbonell, P., A. I. Rodríguez-Escudero, and D. Pujari (2009). "Customer Involvement in New Service Development: An Examination of Antecedents and Outcomes." In: *Journal of Product Innovation Management* 26.5, pp. 536–550.
- Cassar, G. and J. Craig (2009). "An investigation of hindsight bias in nascent venture activity." In: *Journal of Business Venturing* 24.2, pp. 149–164.
- Chandrashekar, R. (2004). "The influence of redundant comparison prices and other price presentation formats on consumersâ evaluations and purchase intentions." In: *Journal of Retailing* 80.1, pp. 53–66.
- Chen, K.-Y. and C. R. Plott (2002). "Information Aggregation Mechanisms: Concept, Design and Implementation for a sales forecasting problem."
- Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. 1st ed. Boston: Harvard Business Press.
- Chesbrough, H. W. (2004). "Managing open innovation." In: *Research-Technology Management* 47.1, pp. 23–26.
- Christen, M. (2005). "Research Note - Cost Uncertainty Is Bliss: The Effect of Competition on the Acquisition of Cost Information for Pricing New Products." In: *Management Science* 51.4, pp. 668–676.
- Christensen, C. M. and J. L. Bower (1996). "Customer power, strategic investment, and the failure of leading firms." In: *Strategic Management Journal* 17.3, pp. 197–218.
- Christensen, C. M., S. Cook, and T. Hall (2005). "Marketing malpractice: the cause and the cure." In: *Harvard business review* 83.12, p. 74.

Bibliography

- Clark, K. B. (1989). "Project scope and project performance: the effect of parts strategy and supplier involvement on product development." In: *Management science* 35.10, pp. 1247–1263.
- Clark, N. (2008). "Société Générale posts record loss on trading scandal, subprime exposure." In: *International Herald Tribune* 19.
- Clemen, R. T. (1989). "Combining forecasts: A review and annotated bibliography." In: *International Journal of Forecasting* 5, pp. 559–583.
- Clemen, R. T. and R. L. Winkler (1993). "Aggregating point estimates: A flexible modeling approach." In: *Management Science* 39.4, pp. 501–515.
- Cooper, R. G. (1988). *Winning at new products*. Kogan Page London.
- Cooper, R. G. (1990). "Stage-Gate Systems : A New Tool for Managing New Products." In: *Business Horizons* 33.3, pp. 44–54.
- Cooper, R. G. (1994). "Perspective third-generation new product processes." In: *Journal of Product Innovation Management* 11.1, pp. 3–14.
- Cooper, R. G., S. J. Edgett, and E. J. Kleinschmidt (1998). "Best practices for managing R&D portfolios." In: *Research-Technology Management* 41.4, pp. 20–33.
- Courtney, H, J Kirkland, and P Viguerie (1997). "Strategy under uncertainty." In: *Harvard Business Review* 6, pp. 67–79.
- Cowgill, B., J. Wolfers, and E. Zitzewitz (2008). "Using Prediction Markets to Track Information Flows: Evidence from Google."
- Creese, R. C. and T. Moore (1990). "Cost Modelling for Concurrent Engineering." In: *Cost Engineering* 32.6, pp. 23–27.
- Dahan, E., A. Soukhoroukova, and M. Spann (2010). "New Product Development 2.0: Preference Markets - How Scalable Securities Markets Identify Winning Product Concepts & Attributes." In: *Journal of Product Innovation Management* 27.7, pp. 937–954.
- Dahan, E., A.J. Kim, and A.W. Lo (2011). "Securities trading of concepts (STOC)." In: *Journal of Marketing Research* 48.6, pp. 497–517.
- Dahlander, L. and D. M. Gann (2010). "How open is innovation?" In: *Research Policy* 39.6, pp. 699–709.
- Dalkey, N. and O. Helmer (1963). "An experimental application of the Delphi method to the use of experts." In: *Management Science* 9.3, pp. 458–467.
- Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1997). "A Theory of Overconfidence, Self-Attribution, and Security Market Under- and Overreactions."

Bibliography

- Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998). "Investor Psychology and Security Market Under- and Overreactions." In: *Journal of Finance* 53.6, pp. 1839–1885.
- DeSanctis, G. and R. B. Gallupe (1987). "A Foundation for the Study of Group Decision Support Systems." In: *Management Science* 33.5, pp. 589–609.
- Deaves, R., E. Luders, and G. Y. Luo (2009). "An Experimental Test of the Impact of Overconfidence and Gender on Trading Activity." In: *Review of Finance* 13.3, pp. 555–575.
- Deck, C., S. Lin, and D. Porter (2013). "Affecting policy by manipulating prediction markets: Experimental evidence." In: *Journal of Economic Behavior & Organization* 85.1, pp. 48–62.
- Denrell, J. and C. Fang (2010). "Predicting the Next Big Thing: Success as a Signal of Poor Judgment." In: *Management Science* 56.10, pp. 1653–1667.
- Duffy, J. (2006). "Agent-based models and human subject experiments." In: *Handbook of Computational Economics*. Vol. 2. 05, pp. 949–1010.
- Dushnitsky, G. (2009). "Entrepreneurial Optimism in the Market for Technological Inventions." In: *Organization Science* 21.1, pp. 150–167.
- Dye, R. (2008). "The Promise of Prediction Markets: A Roundtable." In: *The McKinsey Quarterly*, pp. 83–93.
- Eichner, A. S. (1983). "Why Economics is not yet a Science." In: *Journal of Economic Issues* 17.2, pp. 507–520.
- Eisenhardt, K. M. and B. N. Tabrizi (1995). "Accelerating Adaptive Processes: Product Innovation in the Global Computer Industry." In: *Administrative Science Quarterly* 40.1, p. 84.
- Engel, C., S. Kube, and M. Kurschilgen (2011). "Can We Manage First Impressions in Cooperation Problems? An Experimental Study on 'Broken (and Fixed) Windows'." In: *Journal of Product Innovation Management* 16.1, pp. 52–64.
- Englund, R. L. and R. J. Graham (1999). "From experience: linking projects to strategy." In: *Journal of Product Innovation Management* 16.1, pp. 52–64.
- Enkel, E., C. Kausch, and O. Gassmann (2005). "Managing the Risk of Customer Integration." In: *European Management Journal* 23.2, pp. 203–213.
- Evans, J. (1989). *Bias in human reasoning: Causes and consequences. Essays in cognitive psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., pp. 1–145.
- Evanschitzky, H. et al. (2012). "Success Factors of Product Innovation: An Updated Meta-Analysis." In: *Journal of Product Innovation Management* 29.1, pp. 21–37.

Bibliography

- Fama, E. F. (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work." In: *Journal of Finance* 25.2, pp. 383–417.
- Fama, E. F. and K. R. French (1992). "The cross-section of expected stock returns." In: *Journal of Finance* 47.2, pp. 427–465.
- Feldman, J. M. and J. G. Lynch (1988). "Self-Generated Validity and Other Effects of Measurement on Belief , Attitude , Intention , and Behavior." In: *Journal of Applied Psychology* 73.3, pp. 421–435.
- Felli, J. C., R. E. Kochevar, and B. R. Gritton (2000). "Rethinking project selection at the Monterey Bay Aquarium." In: *Interfaces* 30.6, pp. 49–63.
- Ferguson, E. and T. Cox (1993). "Exploratory Factor Analysis: A Users' Guide." In: *International Journal of Selection and Assessment* 1.2, pp. 84–94.
- Fischbacher, U. (2007). "z-Tree: Zurich toolbox for ready-made economic experiments." In: *Experimental Economics* 10.2, pp. 171–178.
- Fischhoff, B., P. Slovic, and S. Lichtenstein (1977). "Knowing with certainty: The appropriateness of extreme confidence." In: *Journal of Experimental Psychology* 3.4, pp. 552–564.
- Forsythe, R., F. Nelson, and G. R. Neumann (1992). "Anatomy of an experimental political stock market." In: *The American Economic Review* 82.5, pp. 1142–1161.
- Forsythe, R., T. A. Rietz, and T. W. Ross (1999). "Wishes, expectations and actions: a survey on price formation in election stock markets." In: *Journal of Economic Behavior & Organization* 39.1, pp. 83–110.
- Foster, J. (2010). "Productivity, Creative Destruction and Innovation Policy." In: *Australian Business Foundation* February, pp. 1–26.
- Fowler, J. and D. Johnson (2011). "The evolution of overconfidence." In: *Nature* 477.7364, pp. 317–20.
- Franke, N. (2002). *Realtheorie des Marketing: Gestalt und Erkenntnis*. Vol. 124. Mohr Siebeck.
- Franke, N. and F. Piller (2004). "Value creation by toolkits for user innovation and design: The case of the watch market." In: *Journal of Product Innovation Management* 21.6, pp. 401–415.
- Franke, N., E. Von Hippel, and M. Schreier (2006). "Finding Commercially Attractive User Innovations: A Test of Lead-User Theory." In: *Journal of Product Innovation Management* 23.4, pp. 301–315.
- Frey, D. (1978). "Reactions to success and failure in public and in private conditions." In: *Journal of Experimental Social Psychology* 14.2, pp. 172–179.

Bibliography

- Frishammar, J., H. Floren, and J. Wincent (2011). "Beyond managing uncertainty: insights from studying equivocality in the fuzzy front end of product and process innovation projects." In: *IEEE Transactions on Engineering Management* 58.3, pp. 551–563.
- Füller, J. et al. (2006). "Community based innovation: How to integrate members of virtual communities into new product development." In: *Electronic Commerce Research* 6.1, pp. 57–73.
- Gaissmaier, W. and J. N. Marewski (2011). "Forecasting elections with mere recognition from small, lousy samples: A comparison of collective recognition, wisdom of crowds, and representative polls." In: *Judgment and Decision Making* 6.1, pp. 73–88.
- Galasso, A. and T. S. Simcoe (2011). "CEO Overconfidence and Innovation." In: *Management Science* 57.8, pp. 1469–1484.
- Galbraith, J. (1973). *Designing Complex Organizations*. 1st ed. Reading: Addison-Wesley.
- Garcia, R. and R. Calantone (2002). "A critical look at technological innovation typology and innovativeness terminology: a literature review." In: *Journal of Product Innovation Management* 19.2, pp. 110–132.
- Garthwaite, P. H., J. B. Kadane, and A. O'hagan (2005). "Statistical methods for eliciting probability distributions." In: *Journal of the American Statistical Association* 100.470, pp. 680–701.
- Gassmann, O. (2006b). "Opening up the innovation process: towards an agenda." In: *R&D Management*, pp. 223–228.
- Gassmann, O. (2006a). "Opening up the innovation process: towards an agenda." In: *R&D Management* 36.3, pp. 223–228.
- Gassmann, O., C. Schindellegi, and E. Enkel (2010). "Negative side effects of customer integration." In: *International Journal of Technology Management* 50.1, pp. 43–62.
- Gentner, D. and A. L. Stevens (2014). *Mental Models*. 1st ed. Psychology Press.
- Gervais, S. (2001). "Learning to Be Overconfident." In: *Review of Financial Studies* 14.1, pp. 1–27.
- Gigerenzer, G. and H. Brighton (2009). "Homo Heuristicus: Why Biased Minds Make Better Inferences." In: *Topics in Cognitive Science* 1.1, pp. 107–143.
- Gigerenzer, G. and P. M. Todd (2008). "Rationality the fast and frugal way: introduction." In: *Handbook of Experimental Economics Results* 1, pp. 976–986.
- Gilovich, T. and D. Griffin (2002). "Introduction-heuristics and biases: Then and now." In: *Heuristics and Biases: The Psychology of Intuitive Judgment*, pp. 1–18.

Bibliography

- Glaser, E. M., H. H. Abelson, and K. N. Garrison (1983). *Putting knowledge to use: Facilitating the diffusion of knowledge and the implementation of planned change*. San Francisco: Jossey-Bass.
- Glaser, M., M. Nöth, and M. Weber (2003). "Behavioral Finance."
- Gold, J., A. Sekuler, and P. Bennett (2004). "Characterizing perceptual learning with external noise." In: *Cognitive Science* 28.2, pp. 167–207.
- Goldstein, W. M. and R. M. Hogarth (1997). "Judgment and decision research: Some historical context." In: *Research on Judgment and Decision Making: Currents, Connections, and Controversies*, pp. 3–65.
- Goodman, C. M. (1987). "The Delphi technique: a critique." In: *Journal of Advanced Nursing* 12.6, pp. 729–34.
- Gort, C., M. Wang, and M. Siegrist (2008). "Are Pension Fund Managers Overconfident?" In: *Journal of Behavioral Finance* 9.3, pp. 163–170.
- Grabher, G., O. Ibert, and S. Flohr (2008). "The neglected king: The customer in the new knowledge ecology of innovation." In: *Economic Geography* 84.3, pp. 253–280.
- Gracia, E. and J. Herrero (2008). "Is it considered violence? The acceptability of physical punishment of children in Europe." In: *Journal of Marriage and Family* 70. February, pp. 210–217.
- Graefe, A. (2009). "Prediction Markets versus Alternative Methods: Empirical Tests of Accuracy and Acceptability." PhD thesis.
- Graefe, A. and J. S. Armstrong (2011). "Comparing face-to-face meetings, nominal groups, Delphi and prediction markets on an estimation task." In: *International Journal of Forecasting* 27.1, pp. 183–195.
- Granitz, N. A. and J. C. Ward (1996). "Virtual community: A sociocognitive analysis." In: *Advances in Consumer Research* 23, pp. 161–166.
- Green, K. C., J. S. Armstrong, and A. Graefe (2007). "Methods to Elicit Forecasts from Groups: Delphi and Prediction Markets Compared."
- Greenberg, J et al. (1992). "Why do people need self-esteem? Converging evidence that self-esteem serves an anxiety-buffering function." In: *Journal of Personality and Social Psychology* 63.6, pp. 913–22.
- Griffin, A. (1997). "The effect of project and process characteristics on product development cycle time." In: *Journal of Marketing Research* 34.2.
- Gruenfeld, D. H. et al. (1996). "Group composition and decision making: How member familiarity and information distribution affect process and performance." In: *Organizational Behavior and Human Decision Processes* 67.1, pp. 1–15.

Bibliography

- Gupta, A. K., P. E. Tesluk, and M. S. Taylor (2007). "Innovation At and Across Multiple Levels of Analysis." In: *Organization Science* 18.6, pp. 885–897.
- Hair, J. F. et al. (2010). *Multivariate data analysis*. 7th ed. Vol. 7. Prentice Hall Upper Saddle River, NJ.
- Hall, J. et al. (2011). "Managing technological and social uncertainties of innovation: The evolution of Brazilian energy and agriculture." In: *Technological Forecasting and Social Change* 78.7, pp. 1147–1157.
- Hammond, J. S. and R. L. Keeney (1998). "The hidden traps in decision making." In: *Harvard Business Review*.
- Hanft, R. S. and S. P. Korper (1981). "Some notes on uncertainty; Federal policy and innovation." In: *Computers in Biology and Medicine* 11, pp. 1–7.
- Hanson, R. (2003). "Combinatorial information market design." In: *Information Systems Frontiers* 5.1, pp. 107–119.
- Hanson, R., R. Oprea, and D. Porter (2006). "Information aggregation and manipulation in an experimental market." In: *Journal of Economic Behavior & Organization* 60.4, pp. 449–459.
- Hanson, R. et al. (2007). "Can manipulators mislead market observers."
- Harris, E. and R. Woolley (2009). "Facilitating Innovation Through Cognitive Mapping of Uncertainty." In: *International Studies of Management and Organization* 39.1, pp. 70–100.
- Hauser, J., A. Griffin, and G. J. Tellis (2006). "Research on Innovation: A Review and Agenda for Marketing Science." In: *Marketing Science* 25.6, pp. 687–717.
- Hauschildt, J. and S. Salomo (2007). *Innovationsmanagement*. 4th ed. München: Vahlen.
- Hayek, F. A. (1945). "The Use of Knowledge in Society." In: *The American Economic Review* 35.4, pp. 519–530.
- Hayward, M. L., V. P. Rindova, and T. G. Pollock (2004). "Believing one's own press: the causes and consequences of CEO celebrity." In: *Strategic Management Journal* 25.7, pp. 637–653.
- Hayward, M. L., A. Shepherd, and D. Griffin (2006). "A hubris theory of entrepreneurship." In: *Management Science* 52.2, p. 160.
- Healy, P. J. et al. (2010). "Prediction Markets: Alternative Mechanisms for Complex Environments with Few Traders." In: *Management Science* 56.11, pp. 1977–1996.
- Heckhausen, J. and H. Heckhausen (2006). *Motivation und Handeln*. 4th ed. Heidelberg: Springer.

Bibliography

- Heiskanen, E. et al. (2007). "User involvement in radical innovation: are consumers conservative?" In: *European Journal of Innovation Management* 10.4, pp. 489–509.
- Herstatt, C. and E. von Hippel (1992). "From experience: Developing new product concepts via the lead user method: A case study in a low-Tech field." In: *Journal of Product Innovation Management* 9.3, pp. 213–221.
- Herstatt, C., B. Verworn, and A. Nagahira (2004). "Reducing project related uncertainty in the "fuzzy front end" of innovation: a comparison of German and Japanese product innovation projects." In: *International Journal of Product Development* 1.1, pp. 43–65.
- Hilbert, M. (2012). "Toward a synthesis of cognitive biases: how noisy information processing can bias human decision making." In: *Psychological Bulletin* 138.2, pp. 211–37.
- Hiltz, S. R., K. Johnson, and M. Turoff (1986). "Experiments in Group Decision Making: Communication Process and Outcome in Face-to-Face Versus Computerized Conferences." In: *Human Communication Research* 13.2, pp. 225–252.
- Hise, R. T. and L. O'Neal (1990). "Marketing/R&D interaction in new product development: implications for new product success rates." In: *Journal of Product Innovation Management* 7, pp. 142–155.
- Hoetker, G. (2005). "How much you know versus how well I know you: selecting a supplier for a technically innovative component." In: *Strategic Management Journal* 26.1, pp. 75–96.
- Hox, J. (1998). "Multilevel modeling: When and why." In: *Classification, data analysis, and data highways*. Springer, pp. 147–154.
- Hurst, P. (1982). "Ideas into action development and the acceptance of innovations." In: *International Journal of Educational Development* 1.3, pp. 79–102.
- Islei, G. et al. (1991). "A decision support system using judgmental modeling: a case of R&D in the pharmaceutical industry." In: *IEEE Transactions on Engineering Management* 38.3, pp. 202–209.
- Jalonen, H. (2011). "The uncertainty of innovation: a systematic review of the literature." In: *Journal of Management Research* 4.1, pp. 1–47.
- Jeppesen, L. B. and L. Frederiksen (2006). "Why Do Users Contribute to Firm-Hosted User Communities? The Case of Computer-Controlled Music Instruments." In: *Organization Science* 17.1, pp. 45–63.
- Jian, L. and R. Sami (2011). "Aggregation and Manipulation in Prediction Markets: Effects of Trading Mechanism and Information Distribution." In: *Management Science* 58.1, pp. 123–140.

Bibliography

- Johnsen, T. E. (2009). "Supplier involvement in new product development and innovation: Taking stock and looking to the future." In: *Journal of Purchasing and Supply Management* 15.3, pp. 187–197.
- Judd, C. M., D. A. Kenny, and G. H. McClelland (2001). "Estimating and testing mediation and moderation in within-subject designs." In: *Psychological Methods* 6.2, pp. 115–134.
- Kahneman, D. (2003). "Maps of bounded rationality: Psychology for behavioral economics." In: *American economic review* 93.5, pp. 1449–1475.
- Kahneman, D. and A. Tversky (1979). "Prospect theory: An analysis of decision under risk." In: *Econometrica: Journal of the Econometric Society* 47.2, pp. 263–292.
- Kaplan, T. R., I. Luski, and D. Wettstein (2003). "Innovative activity and sunk cost." In: *International Journal of Industrial Management*.
- Karniouchina, E. V. (2011). "Are Virtual Markets Efficient Predictors of New Product Success? The Case of the Hollywood Stock Exchange." In: *Journal of Product Innovation Management* 28.4, pp. 470–484.
- Kashy, D. A. and D. A. Kenny (2000). "The analysis of data from dyads and groups." In: *Handbook of Research Methods in Social and Personality Psychology*. Ed. by H. T. Reis and C. M. Judd. Cambridge: Cambridge University Press, pp. 451–477.
- Katila, R. and G. Ahuja (2002). "Something old, something new: A longitudinal study of search behavior and new product introduction." In: *Academy of management journal* 45.6, pp. 1183–1194.
- Keiser, T. M. and M. R. Burns (1998). *Computer-implemented securities trading system with a virtual specialist function - Pa.No.: US 6,505,174 B1*. Santa Monica.
- Khurana, A. and S. R. Rosenthal (1998). "Towards holistic âfront endsâ in new product development." In: *Journal of Product Innovation Management* 15.1, pp. 57–74.
- Kim, J. and D. Wilemon (2002). "Strategic issues in managing innovationâs fuzzy front-end." In: *European Journal of Innovation Management* 5.1, pp. 27–39.
- Klevatorick, A. K. et al. (1995). "On the sources and significance of interindustry differences in technological opportunities." In: *Research Policy* 24.2, pp. 185–205.
- Klingert, F. (2013). "Corporate prediction markets: A multi-agent simulation based on laboratory experiments." PhD thesis. Hamburg University of Technology.
- Knudsen, M. P. (2007). "The Relative Importance of Interfirm Relationships and Knowledge Transfer for New Product Development Success*." In: *Journal of Product Innovation Management* 24.May, pp. 117–138.

Bibliography

- Koellinger, P., M. Minniti, and C. Schade (2007). "I think I can, I think I can: Overconfidence and entrepreneurial behavior." In: *Journal of Economic Psychology* 28.4, pp. 502–527.
- Koen, P. et al. (2001). "Providing clarity and a common language to the" fuzzy front end"." In: *Research-Technology Management* 44.2, pp. 46–55.
- Kopelman, S., A. S. Rosette, and L. Thompson (2006). "The three faces of Eve: Strategic displays of positive, negative, and neutral emotions in negotiations." In: *Organizational Behavior and Human Decision Processes* 99.1, pp. 81–101.
- Kremer, M., B. Moritz, and E. Siemsen (2011). "Demand forecasting behavior: System neglect and change detection." In: *Management Science* 57.10, pp. 1827–1843.
- Kristensson, P., A. Gustafsson, and T. Archer (2004). "Harnessing the Creative Potential among Users." In: *Journal of Product Innovation Management* 21.1, pp. 4–14.
- Kumar, N., L. W. Stern, and J. C. Anderson (1993). "Conducting Interorganizational Research Using Key Informants." In: *Academy of Management Journal* 36.6, pp. 1633–1651.
- Kunda, Z. (1990). "The case for motivated reasoning." In: *Psychological Bulletin* 108.3, pp. 480–498.
- Lacomb, C., J. Barnett, and Q. Pan (2007). "The imagination market." In: *Information Systems Frontiers* 9.2-3, pp. 245–256.
- Larrick, R. P., K. A. Burson, and J. B. Soll (2007). "Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not)." In: *Organizational Behavior and Human Decision Processes* 102.1, pp. 76–94.
- Laursen, K. and A. Salter (2004). "Searching high and low: what types of firms use universities as a source of innovation?" In: *Research Policy* 33.8, pp. 1201–1215.
- Laursen, K. and A. Salter (2006). "Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms." In: *Strategic Management Journal* 27.2, pp. 131–150.
- Ledyard, J., R. Hanson, and T. Ishikida (2009). "An experimental test of combinatorial information markets." In: *Journal of Economic Behavior & Organization* 69.2, pp. 182–189.
- Lee, J. and F. Veloso (2006). "Inter-Firm Innovation under Uncertainty: Empirical Evidence for Strategic Knowledge-Partitioning."
- Leifer, R., G. C. O. Connor, and M. Rice (2001). "Implementing radical innovation in mature firms : The role of hubs." In: *Academy of Management Executive* 15.3, pp. 102–113.

Bibliography

- Leiponen, A. and C. E. Helfat (2010). "Innovation objectives, knowledge sources, and the benefits of breadth." In: *Strategic Management Journal* 31, pp. 224–236.
- Levitt, S. D. and J. A. List (2007). "What do laboratory experiments measuring social preferences reveal about the real world?" In: *The Journal of Economic Perspectives* 21.2, pp. 153–174.
- Liyanarachchi, G. A. (2007). "Feasibility of using student subjects in accounting experiments: a review." In: *Pacific Accounting Review* 19.1, pp. 47–67.
- Loch, C. H. et al. (2001). "Selecting R&D projects at BMW: A case study of adopting mathematical programming models." In: *IEEE Transactions on Engineering Management* 48.1, pp. 70 –80.
- Loch, CH, ME Solt, and EM Bailey (2008). "Diagnosing Unforeseeable Uncertainty in a New Venture*." In: *Journal of Product Innovation Management* 25, pp. 28–46.
- Lovall, D. and D. Kahneman (2003). "Delusions of Success: How Optimism Undermines Executives' Decisions." In: *Harvard Business Review* April, pp. 1–9.
- Lowe, R. A. and A. A. Ziedonis (2006). "Overoptimism and the performance of entrepreneurial firms." In: *Management Science* 52.2, pp. 173–186.
- Luckner, S. (2007). "A Field Experiment on Monetary Incentives in Prediction Markets." In: *Tagungsband zum Doctoral Consortium der WI 2007*.
- Luckner, S. (2008). "Predictive Power of Markets." PhD thesis.
- Lüthje, C. (2000). "Kundenorientierung im Innovationsprozess." PhD thesis. Wiesbaden.
- Lüthje, C. (2004). "Characteristics of Innovating Users in a Consumer Goods Field: An Empirical Study of Sport-Related Product Consumers." In: *Technovation* 24.9, pp. 683–695.
- Macmillan, N. A. (2002). "Signal detection theory." In: *Stevens' Handbook of Experimental Psychology*.
- Mahajan, J. (1992). "The Overconfidence Effect in Marketing Management Predictions." In: *Journal of Marketing Research* 29.3, p. 329.
- Malmendier, U., G. Tate, and J. Yan (2011). "Overconfidence and Early-Life Experiences: The Effect of Managerial Traits on Corporate Financial Policies." In: *Journal of Finance* 66.5, pp. 1687–1733.
- Marn, M. V., E. V. Roegner, and C. C. Zawada (2003). "Pricing new products." In: *The McKinsey Quarterly* 3.
- Massey, C. and G. Wu (2005). "Detecting Regime Shifts: The Causes of Under- and Overreaction." In: *Management Science* 51.6, pp. 932–947.

Bibliography

- Meisel, J. B. and T. S. Sullivan (2002). "The impact of the Internet on the law and economics of the music industry." In: *info* 4.2, pp. 16–22.
- Miles, J. and M. Shevlin (2001). *Applying regression and correlation: A guide for students and researchers*. Sage.
- Milgrom, P. and N. Stokey (1982). "Information, trade and common knowledge." In: *Journal of Economic Theory* 26.1, pp. 17–27.
- Miljkovic, D. (2005). "Rational Choice and irrational individuals or simply an irrational theory: A critical review of the hypothesis of perfect rationality." In: *The Journal of Socio-Economics* 34.5, pp. 621–634.
- Miller, K. D. and W.-R. Chen (2004). "Variable Organizational Risk Preferences: Tests of the March-Shapira Model." In: *Academy of Management Journal* 47.1, pp. 105–115.
- Miyoshi, T. and M. Azuma (1993). "An empirical study of evaluating software development environment quality." In: *IEEE Transactions on Software Engineering* 19.5, pp. 425–435.
- Mobius, M. M. et al. (2011). "Managing self-confidence: Theory and experimental evidence." Boston.
- Moenaert, R. K. and W. E. Souder (1996). "Context and Antecedents of Information Utility at the R&D / Marketing Interface." In: *Management Science* 42.11, pp. 1592–1610.
- Montoya-Weiss, M. M. and T. M. O'Driscoll (2000). "From Experience: Applying Performance Support Technology in the Fuzzy Front End." In: *Journal of Product Innovation Management* 17.2, pp. 143–161.
- Moore, D. A. and P. J. Healy (2008). "The trouble with overconfidence." In: *Psychological review* 115.2, pp. 502–17.
- Moore, D. A. and D. A. Small (2007). "Error and bias in comparative judgment: on being both better and worse than we think we are." In: *Journal of Personality and Social Psychology* 92.6, pp. 972–89.
- Moore, D. A., J. M. Oesch, and C. Zietsma (2007). "What Competition? Myopic Self-Focus in Market-Entry Decisions." In: *Organization Science* 18.3, pp. 440–454.
- Mullins, J. W. and D. J. Sutherland (1998). "New product development in rapidly changing markets: An exploratory study." In: *Journal of Product Innovation Management* 15, pp. 224–236.
- Mussweiler, T. and B. Englich (2005). "Subliminal anchoring: Judgmental consequences and underlying mechanisms." In: *Organizational Behavior and Human Decision Processes* 98.2, pp. 133–143.

Bibliography

- Neff, J. (2005). "Small ball: Marketers rely on line extensions." In: *Advertising Age (Midwest Region Edition)* 76.15, pp. 10–11.
- Nickerson, R. S. (1998). "Confirmation bias: A ubiquitous phenomenon in many guises." In: *Review of General Psychology* 2.2, pp. 175–220.
- Nisbett, R. E. and L. Ross (1980). *Human inference: Strategies and shortcomings of social judgment*. 1st ed. New Jersey: Prentice Hall, pp. 1–334.
- Nöth, M and Martin Weber (2003). "Information Aggregation with Random Ordering: Cascades and Overconfidence*." In: *The Economic Journal* 113.1995, pp. 166–189.
- Numata, S. et al. (2010). "Research Note: Medical device development in crisis: A movement for technology innovation in health and medicine in Japan." In: *Innovation: Management, Policy & Practice* 12.3, pp. 330–336.
- Oberlechner, T. and C. Osler (2012). "Survival of Overconfidence in Currency Markets." In: *Journal of Financial and Quantitative Analysis*, pp. 1–46.
- Odean, T. (1998). "Volume, Volatility, Price, and Profit When All Traders Are Above Average." In: *Journal of Finance* 53.6, pp. 1887–1934.
- Oliven, K. and T. A. Rietz (2004). "Suckers Are Born but Markets Are Made: Individual Rationality, Arbitrage, and Market Efficiency on an Electronic Futures Market." In: *Management Science* 50.3, pp. 336–351.
- Ortner, G. (1998). "Forecasting markets—An industrial application."
- Ostrover, S. (2005). "Employing information markets to achieve truly collaborative sales forecasting." In: *Journal of Business Forecasting* 24.1, pp. 9–12.
- Othman, A. and T. Sandholm (2010). "Automated market-making in the large: the Gates Hillman prediction market." In: *Proceedings of the 11th ACM conference on Electronic commerce*. ACM, pp. 367–376.
- Ottaviani, M. (2009). "The Design of Idea Markets : An Economist's Perspective." In: *Journal of Prediction Markets* 3.1, pp. 41–44.
- Ozer, M. (2005). "Factors which influence decision making in new product evaluation." In: *European Journal of Operational Research* 163.3, pp. 784–801.
- Pearlstein, S. (2003). *Misplacing Trust in the Markets*.
- Peer, E. (2012). "Professionally biased: Evidence for misestimations of driving speed, journey time and time-savings among taxi and car drivers." In: *Judgment and Decision Making* 7.2, pp. 165–172.
- Persky, J. (1995). "Retrospectives The Ethology of Homo Economicus." In: 9.2, pp. 221–231.

Bibliography

- Peterson, R. A. (1994). "A meta-analysis of Cronbach's coefficient alpha." In: *Journal of consumer research* 21.9, pp. 381–391.
- Piller, F. T. and D. Walcher (2006). "Toolkits for idea competitions: a novel method to integrate users in new product development." In: *R&D Management* 36.3, pp. 307–318.
- Plott, C. R. and S. Sunder (1982). "Efficiency of Experimental Security Markets with Insider Information: an Application of Rational-Expectations Models." In: *The Journal of Political Economy* 90.4, pp. 663–698.
- Plott, C. R. and S. Sunder (1988). "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets." In: *Journal of the Econometric Society* 56.5, pp. 1085–1118.
- Poetz, M. K. and M. Schreier (2012). "The Value of Crowdsourcing: Can Users Really Compete with Professionals in Generating New Product Ideas?" In: *Journal of Product Innovation Management* 29.2, pp. 245–256.
- Pyszczynski, T. and J. Greenberg (1987). "Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model." In: *Advances in Experimental Social Psychology* 20, pp. 297–339.
- Pyszczynski, T., J. Greenberg, and K. Holt (1985). "Maintaining Consistency between Self-Serving Beliefs and Available Data: A Bias in Information Evaluation." In: *Personality and Social Psychology Bulletin* 11.2, pp. 179–190.
- Radzevick, J. R. and D. A. Moore (2011). "Competing to Be Certain (But Wrong): Market Dynamics and Excessive Confidence in Judgment." In: *Management Science* 57.1, pp. 93–106.
- Ragatz, G. L., R. B. Handfield, and K. J. Petersen (2002). "Benefits associated with supplier integration into new product development under conditions of technology uncertainty." In: *Journal of Business Research* 55.5, pp. 389–400.
- Raudenbush, S. W. and A. S. Byrk (2002). *Hierarchical Linear Models - Applications and Data Analysis Methods*. 2nd ed. Thousand Oaks: Sage, pp. 1–483.
- Reid, S. E. and U. De Brentani (2004). "The Fuzzy Front End of New Product Development for Discontinuous Innovations: A Theoretical Model." In: *Journal of Product Innovation Management* 21.3, pp. 170–184.
- Remus, W (1989). "Using students as subjects in experiments on decision support systems." In: *Proceedings of the Twenty-Second Annual Hawaii International Conference on System Sciences*. Vol. 3. 1, pp. 176–180.
- Rheinberg, F. (1997). "Motivation." In: *Stuttgart: Kohlhammer*.

Bibliography

- Rhode, P. and K. Strumpf (2004). "Historical presidential betting markets." In: *The Journal of Economic Perspectives* 18.2, pp. 127–142.
- Rindfleisch, A. and C. Moorman (2001). "The acquisition and utilization of information in new product alliances: A strength-of-ties perspective." In: *Journal of Marketing* 65.April, pp. 1–18.
- Ritov, I. and C. J. Brainerd (1992). "Status-quo and omission biases." In: *Journal of Risk and Uncertainty* 5.1, pp. 49–61.
- Roberts, E. B. (2007). "Managing invention and innovation." In: *Research-Technology Management* 50.1, pp. 35–54.
- Rosenbloom, E. S. and W. Notz (2006). "Statistical Tests of Real-Money versus Play-Money Prediction Markets." In: *Electronic Markets* 16.1, pp. 63–69.
- Roth, P. L. (1994). "Missing data: A conceptual review for applied psychologists." In: *Personnel Psychology* 47.3, pp. 537–560.
- Rothaermel, F. T. and A. M. Hess (2007). "Building Dynamic Capabilities: Innovation Driven by Individual-, Firm-, and Network-Level Effects." In: *Organization Science* 18.6, pp. 898–921.
- Rothermund, K. and A. Eder (2011). *Motivation und Emotion*. 1st ed. Wiesbaden: VS Verlag.
- Rousseau, D. M. (1985). "Issues of level in organizational research: Multi-level and cross-level perspectives." In: *Research in Organizational Behavior* 7.1, pp. 1–37.
- Rowe, G. (1992). "Perspectives on Expertise in the Aggregation of Judgments." In: *Expertise and Decision Support*. Springer, pp. 155–180.
- Rowe, G. and G. Wright (2011). "The Delphi technique: Past, present, and future prospects—Introduction to the special issue." In: *Technological Forecasting and Social Change* 78.9, pp. 1487–1490.
- Ruttan, V. W. (1959). "Usher and Schumpeter on invention, innovation, and technological change." In: *The Quarterly Journal of Economics* 73.4, pp. 596–606.
- Samuelson, W. and R. Zeckhauser (1988). "Status quo bias in decision making." In: *Journal of risk and uncertainty* 59, pp. 7–59.
- Scandura, T. A. and E. A. Williams (2000). "Research Methodology in Management: Current Practices, Trends, and Implications for Future Research." In: *Academy of Management Journal* 43.6, pp. 1248–1264.
- Scheibe, M., M. Skutsch, and J. Schofer (1975). "Experiments in Delphi methodology." In: *The Delphi Method: Techniques and Applications*, pp. 262–287.

Bibliography

- Scheibehenne, B. and A. Bröder (2007). "Predicting Wimbledon 2005 tennis results by mere player name recognition." In: *International Journal of Forecasting* 23.3, pp. 415–426.
- Schmidt, J. B., K. R. Sarangee, and M. M. Montoya (2009). "Exploring New Product Development Project Review Practices." In: *Journal of Product Innovation Management* 26.5, pp. 520–535.
- Schrader, S., W. M. Riggs, and R. P. Smith (1993). "Choice over uncertainty and ambiguity in technical problem solving." In: *Journal of Engineering and Technology Management* 10.1-2, pp. 73–99.
- Schrader, U. and T. Hennig-Thurau (2009). "VHB-JOURQUAL2 : Method, Results, and Implications of the German Academic Association for Business Research's Journal Ranking." In: *BuR - Business Research* 2.2, pp. 180–204.
- Schumpeter, J. A. (1939). *Business Cycles - Volume I*. 1st Editio. Mansfield Centre: Martino Publishing, pp. 1–448.
- Schütz, P. and B. Bloch (2006). "The âsilo-virusâ: diagnosing and curing departmental groupthink." In: *Team Performance Management* 12.1/2, pp. 31–43.
- Schweisfurth, T. G. (2012). "Lead users inside the boundaries of the firm: Embedded lead users." PhD thesis.
- Sent, E. M. (2004). "Behavioral economics: How psychology made its (limited) way back into economics." In: *History of Political Economy* 36.4, pp. 735–760.
- Serfas, S. (2011). "Cognitive Biases in the Capital Investment Context." PhD thesis.
- Servan-Schreiber, E. et al. (2004). "Prediction markets: Does money matter?" In: *Electronic Markets* 14.3, pp. 243–251.
- Seybert, N. and R. Bloomfield (2009). "Contagion of Wishful Thinking in Markets." In: *Management Science* 55.5, pp. 738–751.
- Shane, S. (1995). "Uncertainty avoidance and the preference for innovation championing roles." In: *Journal of International Business Studies*, pp. 47–68.
- Shaughnessey, J. J., E. B. Zechmeister, and J. S. Zechmeister (2012). *Research Methods in Psychology*. 9th ed. New York: McGraw-Hill.
- Shepperd, J., W. Malone, and K. Sweeny (2008). "Exploring Causes of the Self-serving Bias." In: *Social and Personality Psychology Compass* 2.2, pp. 895–908.
- Simon, H. A. (1955). "A behavioral model of rational choice." In: *The Quarterly Journal of Economics* 69.2, pp. 99–118.
- Simon, H. A. (1959). "Theories of decision-making in economics and behavioral science." In: *The American Economic Review* 49.3, pp. 253–283.

Bibliography

- Simon, M. and S.M. Houghton (2003). "The Relationship between Overconfidence and the Introduction of Risky Products: Evidence from a Field Study." In: *Academy of Management Journal* 46.2, pp. 139–150.
- Simon, M. and R. C. Shrader (2012). "Entrepreneurial actions and optimistic overconfidence: The role of motivated reasoning in new product introductions." In: *Journal of Business Venturing* 27.3, pp. 291–309.
- Slamka, C. and W. Jank (2009). "Second Generation Prediction Markets for Information Aggregation: A Comparison of Payoff Mechanisms." In: *International Journal of Forecasting*, pp. 1–33.
- Smith, P. G. and D. G. Reinertsen (1992). "Shortening the product development cycle." In: *Research-Technology Management* May-June, pp. 44–49.
- Smith, V. L. (1962). "An experimental study of competitive market behavior." In: *The Journal of Political Economy* 70.2, pp. 111–137.
- Snizek, J. A. and R. A. Henry (1989). "Accuracy and confidence in group judgment." In: *Organizational Behavior and Human Decision Processes* 43.1, pp. 1–28.
- Snijders, T. A. B. and R. J. Bosker (2012). *Multilevel Analysis*. 2nd ed. London: SAGE Publications, pp. 1–354.
- Soll, J. B. and J. Klayman (2004). "Overconfidence in interval estimates." In: *Journal of Experimental Psychology* 30.2, pp. 299–314.
- Sonnemann, U. (2008). "Do individual judgment biases diminish in markets ? The case of partition-dependence in prediction markets." PhD thesis.
- Sorensen, F., J. Mattsson, and J. Sundbo (2010). "Experimental methods in innovation research." In: *Research Policy* 39.3, pp. 313–322.
- Sorescu, A. B. and J. Spanjol (2008). "Innovation's Effect on Firm Value and Risk : Insights from Consumer." In: *Journal of Marketing* 72.March, pp. 114–132.
- Souder, W. E. (1988). "Managing relations between R&D and marketing in new product development projects." In: *Journal of Product Innovation Management* 5.1, pp. 6–19.
- Souder, W. E. and A. K. Chakrabarti (1978). "The R&D/Marketing Interface: Results from an Empirical Study of Innovation Projects." In: *IEEE Transactions on Engineering Management* 25.5, pp. 88–93.
- Souder, W. E. and R. K. Moenaert (1992). "Integrating Marketing and R&D Project Personnel Within Innovation Projects: an Information Uncertainty Model." In: *Journal of Management Studies* 29.4, pp. 485–512.
- Soukhoroukova, A. (2006). "New product development with Internet-based information markets: Theory and empirical application." In: *ECIS2006*.

Bibliography

- Soukhoroukova, A. (2007). "Produktinnovation mit Informationsmärkten." PhD thesis.
- Soukhoroukova, A., M. Spann, and B. Skiera (2012). "Sourcing , Filtering , and Evaluating New Product Ideas : An Empirical Exploration of the Performance of Idea Markets." In: *Journal of Product Innovation Management* 29.1, pp. 100–112.
- Spann, M. (2002). "Virtuelle Börsen als Instrument zur Marktforschung." PhD thesis.
- Spann, M. and B. Skiera (2003a). "Internet-Based Virtual Stock Markets for Business Forecasting." In: *Management Science* 49.10, pp. 1310–1326.
- Spann, M. and B. Skiera (2003b). "Taking stock of virtual markets." In: *OR MS TODAY* 30.5, pp. 20–25.
- Spann, M., H. Ernst, and J. H. Soll (2009). "Identification of Lead Users for Consumer Products via Virtual Stock Markets." In: *Journal of Product Innovation Management*.
- Spears, B., C. Lacombe, and J. Barnett (2009). "Examining Trader Behavior In Idea Markets : An Implementation Of GE's Imagination Markets." In: *Journal of Prediction Markets* 3.1, pp. 17–39.
- Stasser, G. and W. Titus (1985). "Pooling of unshared information in group decision making: Biased information sampling during discussion." In: *Journal of Personality and Social Psychology* 48.6, p. 1467.
- Statman, M., S. Thorley, and K. Vorkink (2006). "Investor Overconfidence and Trading Volume." In: *Review of Financial Studies* 19.4, pp. 1531–1565.
- Stevens, G. A. and J. Burley (2003). "Piloting the rocket of radical innovation." In: *Research-Technology Management* 46.2, pp. 16–25.
- Stockstrom, C. and C. Herstatt (2008). "Planning and uncertainty in new product development." In: *R&D Management* 38.5, pp. 480–490.
- Surowiecki, J. (2005). *The wisdom of crowds*. First edit. New York: Anchor.
- Tatikonda, M.V. and S.R. Rosenthal (2000). "Technology novelty, project complexity, and product development project execution success: a deeper look at task uncertainty in product innovation." In: *IEEE Transactions on Engineering Management* 47.1, pp. 74–87.
- Tether, B. S. (2002). "Who co-operates for innovation , and why An empirical analysis." In: *Research Policy* 31.6, pp. 947–967.
- Tether, B. S. and A. Tajar (2008). "Beyond industryâuniversity links: Sourcing knowledge for innovation from consultants, private research organisations and the public science-base." In: *Research Policy* 37.6-7, pp. 1079–1095.

Bibliography

- Tetlock, P. E. and A. Levi (1982). "Attribution bias: On the inconclusiveness of the cognition-motivation debate." In: *Journal of Experimental Social Psychology* 18.1, pp. 68–88.
- Tirole, J. (1988). *The theory of industrial organization*. MIT press.
- Trendence Institut (2011). *Trendence Graduate Barometer 2011 - German Engineering Edition*.
- Trevelyan, R. (2008). "Optimism, overconfidence and entrepreneurial activity." In: *Management Decision* 46.7, pp. 986–1001.
- Trivers, R. (2011). *Deceit and Self-Deception: Fooling Yourself the Better to Fool Others*. New York: Allen Lane, pp. 1–416.
- Tversky, A. and D. Kahneman (1974). "Judgment under Uncertainty: Heuristics and Biases." In: *Science* 185.4157, pp. 1124–31.
- Tversky, A. and D. Kahneman (1975). "Judgment under Uncertainty: Heuristics and Biases." English. In: *Utility, Probability, and Human Decision Making SE - 8*. Ed. by Dirk Wendt and Charles Vlek. Vol. 11. Theory and Decision Library. Springer, pp. 141–162.
- Tversky, A. and D. Kahneman (1991). "Loss aversion in riskless choice: A reference-dependent model." In: *The Quarterly Journal of Economics* 89.
- Ulwick, A. W. (2002). "Turn customers input into innovation." In: *Harvard Business Review*.
- Van Bruggen, G. H., G. L. Lilien, and M. Kacker (2002). "Informants in organizational marketing research: Why use multiple informants and how to aggregate responses." In: *Journal of Marketing Research* 39.4, pp. 469–478.
- Van Bruggen, G. H. et al. (2010). "Prediction Markets as institutional forecasting support systems." In: *Decision Support Systems* 49.4, pp. 404–416.
- Van De, A. and A. L. Delbecq (1971). "Nominal versus interacting group processes for committee decision-making effectiveness." In: *Academy of Management Journal* 14.2, pp. 203–212.
- Van Kleef, G. A. et al. (2010). "On angry leaders and agreeable followers. How leaders' emotions and followers' personalities shape motivation and team performance." In: *Psychological Science* 21.12, pp. 1827–34.
- Van de Vrande, V. et al. (2009). "Open innovation in SMEs: Trends, motives and management challenges." In: *Technovation* 29.6-7, pp. 423–437.

Bibliography

- Vany, A. D. and W. D. Walls (1999). "Uncertainty in the movie industry: Does star power reduce the terror of the box office?" In: *Journal of Cultural Economics* 23.1, pp. 285–319.
- Vermeulen, P., R. Buch, and R. Greenwood (2007). "The Impact of Governmental Policies in Institutional Fields: The Case of Innovation in the Dutch Concrete Industry." In: *Organization Studies* 28.4, pp. 515–540.
- Von Hippel, E. (1986). "Lead Users: A Source Of Novel Product Concepts." In: *Management Science* 32.7, pp. 791–806.
- Von Hippel, E. (2005). *Democratizing Innovation*. Boston: The MIT Press, pp. 1–177.
- Von Hippel, E. (2007). "Horizontal innovation networks - by and for users." In: *Industrial and Corporate Change* 16.2, pp. 293–315.
- Von Hippel, E. and R. Katz (2002). "Shifting Innovation to Users via Toolkits." In: *Management Science* 48.7, pp. 821–833.
- Von Hippel, W. and R. Trivers (2011). "The evolution and psychology of self-deception." In: *Behavioral and Brain Sciences* 34.01, pp. 1–16.
- Von Winterfeldt, D. and W. Edwards (1986). *Decision analysis and behavioral research*. Vol. 1. Cambridge University Press Cambridge.
- Wally, S. and R. J. Baum (1994). "Personal and structural determinants of the pace of strategic decision making." In: *Academy of Management Journal* 37.4, pp. 932–956.
- Weber, E. U., A. R. Blais, and N. E. Betz (2002). "A domain-specific risk attitude scale: Measuring risk perceptions and risk behaviors." In: *Journal of Behavioral Decision Making* 290.8, pp. 263–290.
- Weber, M (2000). "The impact of endowment framing on market prices â an experimental analysis." In: *Journal of Economic Behavior & Organization* 41.2, pp. 159–176.
- Weizsäcker, G (2010). "Do we follow others when we should? A simple test of rational expectations." In: *The American Economic Review* 100.December, pp. 2340–2360.
- Willer, D. and H. A. Walker (2007). *Building experiments: Testing social theory*. Stanford University Press.
- Wolfers, J. and E. Zitzewitz (2004). "Prediction Markets." In: *The Journal of Economic Perspectives* 18.2, pp. 107–126.
- Woudenberg, F. (1991). "An evaluation of Delphi." In: *Technological Forecasting and Social Change* 40.2, pp. 131–150.
- Wu, S.-W., J. E.V. Johnson, and M.-C. Sung (2008). "Overconfidence in Judgments: The Evidence, the Implications and the Limitations." In: *Journal of Prediction Markets* 2.1, pp. 73–90.

Bibliography

- Wynstra, F., A. Van Weele, and M. Weggemann (2001). "Managing supplier involvement in product development:: Three critical issues." In: *European Management Journal* 19.2, pp. 157–167.
- Xu, S., F. Wu, and E. Cavusgil (2013). "Complements or Substitutes? Internal Technological Strength, Competitor Alliance Participation, and Innovation Development." In: *Journal of Product Innovation Management* 30.4.
- Zacharakis, A. L. and D. A. Shepherd (2001). "The nature of information and overconfidence on venture capitalists' decision making." In: *Journal of Business Venturing* 16.4, pp. 311–332.
- Zack, M. H. (2001). "If managing knowledge is the solution, then what's the problem?" In: *Knowledge Management and Business Model Innovation*. Ed. by Yogesh Malhotra. Idea Group Publishing.
- Zhang, Q. and W. J. Doll (2001). "The fuzzy front end and success of new product development: a causal model." In: *European Journal of Innovation Management* 4.2, pp. 95–112.
- Zigurs, I. and B. K. Buckland (1998). "A Theory of Task/Technology Fit and Group Support Systems Effectiveness." In: *MIS quarterly* September.

Part VI.

Appendix

Nr	Author	Year	Journal	Title
1	Ariely, D. et. al.	2005	JMR	When Do Losses Loom Larger Than Gains?
2	Ashford, S. & Tsui, A.	1991	AMJ	Self-regulation for managerial effectiveness: The role of active feedback seeking
3	Astebro, T. & Elhedhli, S.	2006	ManSci	The Effectiveness of Simple Decision Heuristics: Forecasting Commercial Success for Early-Stage Ventures
4	Audia, P. et al.	2000	AMJ	The paradox of success: An archival and a laboratory study of strategic persistence following radical environmental change
5	Barnett, W. & Pontikes, E.	2008	ManSci	The Red Queen, Success Bias, and Organizational Inertia
6	Baron, R.	1998	JBV	Cognitive mechanisms in entrepreneurship: Why and when entrepreneurs think differently than other people
7	Baron, R. et. al.	2012	JBV	Entrepreneurs' dispositional positive affect: The potential benefits – and potential costs – of being “up”
8	Biais, B. & Weber, M.	2009	ManSci	Hindsight Bias, Risk Perception, and Investment Performance
9	Billett, M. & Qian, Y.	2008	ManSci	Are Overconfident CEOs Born or Made? Evidence of Self-Attribution Bias from Frequent Acquirers
10	Blume, B. & Covin, J.	2011	JBV	Attributions to intuition in the venture founding process: Do entrepreneurs actually use intuition or just say that they do?
11	Boeker W.	1997	AMJ	Strategic change: The influence of managerial characteristics and organizational growth
12	Burmeister, K. & Schade, C.	2007	JBV	Are entrepreneurs' decisions more biased? An experimental investigation of the susceptibility to status quo bias
13	Busenitz, L. & Barney, J.	1997	JBV	Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making
14	Camerer, C.	2005	JMR	Three Cheers—Psychological, Theoretical, Empirical—for Loss Aversion
15	Cassar, G. & Craig, J.	2009	JBV	An investigation of hindsight bias in nascent venture activity
16	Cesarini, D.	2012	ManSci	The Behavioral Genetics of Behavioral Anomalies
17	Chen, X. et. al.	2009	AMJ	Entrepreneur passion and preparedness in business plan presentations: a persuasion analysis of venture capitalists' funding decisions
18	Chrisman, J. & Patel, P.	2012	AMJ	Variations in R&D investments of family and nonfamily firms: Behavioral agency and myopic loss aversion perspectives
19	Cooper, A. et. al.	1995	JBV	Entrepreneurial information search
20	Denrell, J. & Fang, C.	2010	ManSci	Predicting the Next Big Thing: Success as a Signal of Poor Judgment

Figure 9.1.: Articles for literature analysis

21	DeTienne, D. et al.	2008	JBV	The fallacy of “only the strong survive”: The effects of extrinsic motivation on the persistence decisions for under-performing firms
22	Dillon, R. & Tinsley, C.	2008	ManSci	How Near-Misses Influence Decision Making Under Risk: A Missed Opportunity for Learning
23	Dushnitsky, G.	2010	OrgSci	Entrepreneurial Optimism in the Market for Technological Inventions
24	Flynn, F. & Wiltermuth, S.	2010	AMJ	Who's with me? False consensus, brokerage, and ethical decision making in organizations
25	Fox, C. & Clemen, R.	2005	ManSci	Subjective Probability Assessment in Decision Analysis: Partition Dependence and Bias Toward the Ignorance Prior
26	Franke, N.	2010	ManSci	The “I Designed It Myself” Effect in Mass Customization
27	Fuchs, C. et al.	2010	JM	The Psychological Effects of Empowerment Strategies on Consumers' Product Demand
28	Galasso, A. & Simcoe, T.	2011	ManSci	CEO Overconfidence and Innovation
29	Gatewood, E. et al.	1995	JBV	A longitudinal study of cognitive factors influencing start-up behaviors and success at venture creation
30	Greve, H.	2003	AMJ	A behavioral theory of R&D expenditures and innovations: Evidence from shipbuilding
31	Haridon, O. et al.	2011	ManSci	Experienced vs. Described Uncertainty: Do We Need Two Prospect Theory Specifications?
33	Hayward, M. et al.	2006	ManSci	A Hubris Theory of Entrepreneurship
34	Hayward, M. et al.	2010	JBV	Beyond hubris: How highly confident entrepreneurs rebound to venture again
35	Hilary, G. & Menzly, L.	2006	ManSci	Does Past Success Lead Analysts to Become Overconfident?
36	Hiller, N. & Hambrick, D.	2005	SMJ	Conceptualizing executive hubris: the role of (hyper-) core self-evaluations in strategic decision-making
37	Hmieleski, K. & Baron, R.	2009	AMJ	Entrepreneurs' optimism and new venture performance: A social cognitive perspective
38	Hogarth, R. & Karelaia, N.	2011	OrgSci	Entrepreneurial success and failure: Confidence and fallible judgment
39	Huy, Q.	2011	SMJ	How middle managers' group-focus emotions and social identities influence strategy implementation
40	Jegers, M.	1991	AMJ	Prospect Theory and the Risk-Return Relation: Some Belgian Evidence
41	Kahneman, D. & Lovallo, D.	1993	ManSci	Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking

Figure 9.2.: Articles for literature analysis continued

42	Karniouchina, E.	2011	JPIM	Are Virtual Markets Efficient Predictors of New Product Success? The Case of the Hollywood Stock Exchange
43	Kremer, M. et al.	2011	ManSci	Demand Forecasting Behavior: System Neglect and Change Detection
44	Li, J. & Tang, Y.	2010	AMJ	CEO Hubris and Firm Risk Taking in China: The Moderating Role of Managerial Discretion
45	Lowe, R. & Ziedonis, A.	2006	ManSci	Overoptimism and the Performance of Entrepreneurial Firms
46	Mahajan, J.	1992	JMR	The Overconfidence Effect in Marketing Management Predictions
47	Miller, K. & Chen, W.	2004	AMJ	Variable Organizational Risk Preferences: Tests of March-Shapira Model
48	Miller, K. & Shapira, Z.	2004	SMJ	An empirical test of heuristics and biases affecting real option valuation
49	Miller, K.	2002	SMJ	Knowledge inventories and managerial myopia
51	Mitchell, R. & Shepherd, D.	2010	JBV	To thine own self be true: Images of self, images of opportunity, and entrepreneurial action
52	Moore, D. et al.	2007	OrgSci	What Competition? Myopic Self-Focus in Market-Entry Decisions
53	Murray, G. & Lott, J.	1995	RP	Have UK venture capitalists a bias against investment in new technology-based firms?
54	Novemsky, N. & Kahneman, D.	2005	JMR	How Do Intentions Affect Loss Aversion?
55	Palich, L. & Bagby, R.	1995	JBV	Using Cognitive Theory to Explain Entrepreneurial Risk-Taking: Challenging Conventional Wisdom
56	Parker, S.	2004	JBV	Learning About the Unknown: How Fast Do Entrepreneurs Adjust Their Beliefs?
57	Radzevick, J. & Moore, D.	2011	ManSci	Competing to Be Certain (But Wrong): Market Dynamics and Excessive Confidence in Judgment
58	Schmidt, J. & Calantone, R.	1998	JPIM	Are Really New Product Development Projects Harder to Shut Down
59	Schweitzer, M. & Cachon, G.	2000	ManSci	Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence
60	Seo, M. et al.	2010	AMJ	Affect and the Framing Effect within Individuals over Time: Risk Taking in a Dynamic Investment Simulation
61	Seo, M. & Feldman Barrett, L.	2007	AMJ	Being emotional during decision making—Good or bad? An empirical investigation

Figure 9.3.: Articles for literature analysis continued

62	Shepherd, D. et al.	2009	JBV	Moving forward: Balancing the financial and emotional costs of business failure
63	Simon, M. et al.	1999	JBV	Cognitive biases, risk perception, and venture formation: How individuals decide to start companies
64	Simon, M. & Shrader, R.	2012	JBV	Entrepreneurial actions and optimistic overconfidence: The role of motivated reasoning in new product introductions
65	Simon, M. & Houghton, S.	2003	AMJ	The relationship between overconfidence and the introduction of risky products: Evidence from a field study
66	Sitkin, S. & Weingart, L	1995	AMJ	Determinants of risky decision-making behavior: A test of the mediating role of risk perceptions and propensity
67	Starbuck, W.	2009	OrgSci	Cognitive Reactions to Rare Events: Perceptions, Uncertainty, and Learning
68	Tinsley, C. et al.	2012	ManSci	How Near-Miss Events Amplify or Attenuate Risky Decision Making
69	Townsend, D. et al.	2010	JBV	To start or not to start: Outcome and ability expectations in the decision to start a new venture
70	Ucbasaran, D.	2010	JBV	The nature of entrepreneurial experience, business failure and comparative optimism
71	Van den Steen, E.	2011	ManSci	Overconfidence by Bayesian-Rational Agents
72	van Bruggen, G. et al.	1998	ManSci	Improving Decision Making by Means of a Marketing Decision Support System
73	Yadav, M. et al.	2007	JM	Managing the Future: CEO Attention and Innovation Outcomes
74	Zhao, M. et al.	2012	JPIM	Imagination Difficulty and New Product Evaluation

Figure 9.4.: Articles for literature analysis continued

Treatment Check Evaluation Questions – Part 1

1. Mini market share
How successful is the Mini compared to VW Polo and Fiat Punto? If the sales of all three models make up for 100%, what would be the Mini's percentage market share in the first quarter of 2011?
2. CO2 Emissions per kilometers of new cars
By how many percent did the average CO2 Emission per kilometers regress among newly registered cars from 2006 to 2011 in Germany?
3. Cars licensed to women
On January 1st 2011, what percentage of cars in Germany was registered women?
4. Small vehicles among the Top Ten
VW Polo and Opel Corsa are among the top ten cars in Germany newly registered in the first quarter of 2011. How many percent of the top ten sales are accounted for by these two vehicles?
5. Growing electric fleet
At least in media, cars with electric drives are prominently featured. By what percentage did the amount of registered electric vehicles increase from 1.1.2010 to 1.1.2011 in Germany?
6. VW Top seller
The top selling Volkswagen models in May 2011 are Polo, Passat, Golf and Tiguan. How many percent of new registrations of these four vehicles are accounted for by the VW Passat?
7. Private registrations
How many percent of cars were registered to private individuals (and not to business entities) in Germany on 1.1.2011?
8. Exports to Europe
German cars are also manufactured for international markets. Of all cars manufactured in Germany in 2010, how many percent were exported in European countries?
9. VW Passat
By how many percent did the Volkswagen Passat registrations in Germany increase in July 2011 compared to the same month in 2010?
10. Italian manufacturers
86.811 cars by the two big Italian manufacturers Fiat and Alfa Romeo were registered in 2010. What percentage of these registrations is accounted for by the more expensive Alfa Romeo brand?

Figure 9.5.: Treatment Check - Evaluation questions

Treatment Check Evaluation Questions – Part 2

1. VW Golf Plus
One of Germany's most famous cars is the Volkswagen Golf. 30 year olds are even referred to as the generation "Golf", as they grew up with the car and it is very popular with them. But what percentage of new Volkswagen Golf Plus were actually bought by individuals older than 60 in the first quarter of 2011 in Germany?
2. Fuel-consumption development
By what percentage did the fuel consumption of newly registered cars in Germany decline from between 1991 and 2011?
3. German manufacturer market share
In the first quarter of 2011, 1.026.000 cars were newly registered in Germany. What percentage of these cars was made by German manufacturers?
4. Production of cars, trucks and buses
In 2009, 5.209.857 cars, trucks and busses were manufactured in Germany. By what percentage has the production volume increased in 2010?
5. Registration growth of the Mercedes SLK
The new SLK was introduced in 2011. By what percentage did the sales figures for the Mercedes SLK increase in the first Quarter in 2011 compared to the first quarter in 2010?
6. Engine power
Average car-engine power in Germany in 1996 was 97.6 HP (PS, or XY KW) . By how many percent did average engine power increase over the last 15 years?
7. Fiat 500 buyer
How many percent of Fiat 500 buyers are female?
8. Car color
Over the last years, an increasing number of the newly registered cars were painted white. What percentage of all newly registered cars in Germany in 2010 was really painted white?
9. Compact-class cars
The class of compact cars (like VW Golf, Ford Focus) is generating the most car revenues in Germany in the second quarter 2011. How many percent of all newly registered cars belong to the compact class?
10. Scrapping bonus ("Abwrackprämie")
In 2009, the German government provided a scrapping bonus when owners decided to scrap their very old cars and buy a brand new car instead. By what percentage did the number of scrapped cars increase compared to the previous year?

Figure 9.6.: Treatment Check - Evaluation questions continued

Instructions for the Experiment

Introduction

This experiment is concerned with a mechanism that can be used to evaluate innovative products: information markets. The experiment will consist of two parts. Both parts offer you the chance to win an Amazon voucher. **2* 20,- Euros in the first part** of the experiment and **12 * 20,- Euros in the second part** of the experiment.

In the first part, we are going to test your estimation performance in a series of domain specific estimation tasks. Afterwards you will be asked for your personal self-assessment.

In the second part, you will compete with other individuals in an information market, where you can generate profits from your estimations in a market environment.

Before the Live – Experiment, you will get to view a detailed instruction video.

Basic and important rules

1. Try to estimate, guess or trade as well as you can. If you deliberately enter false and implausible values, you are lowering your chance of winning and you are making the scientific results useless.
2. If you are lacking information regarding the questions, use your gut feeling.
3. During the experiment, please stay at your booth, leave your headphones on and do not cheat. You would deprive yourself of the chance to win and make the scientific results useless for us.
4. After the experiment, please do not talk to other and later participants about the experiment, the tasks and your answers and performance. Again, you would deprive yourself of the chance to win and make the scientific results useless for us.
5. Thank you for your participation and have fun!

Part 1: Estimation task

- You will answer a set of 10 estimation questions
- The questions will deal with a specific market of well-known products
- After the 10 questions, you will be asked to state how many questions you believe to have answered correctly.
- Among all subjects, we will raffle a **20,-€Amazon** voucher. Your probability to win the voucher will highly depend on two things: 1. Answer as many questions as possible correctly, 2. Try to estimate your own performance as accurately as possible.

Part 2: Information market

In an information market, outcomes of uncertain events are traded by the participants. In our case, the outcomes of questions regarding innovations and new products in a well-known market will be traded.

What are information markets?

Basic idea

- Beliefs regarding future market shares become tradeable
- Participants can trade their beliefs through stocks
- The stock prices reflect the participants' beliefs regarding the future market share

Example: Information market to predict the global market share of the iPhone 5 on January 1st 2013

- Payout per stock = 1 \$ * global market share of iPhone 5 on January 1st 2013
- Expected market share for iPhone 5 on given date: 50% -> payout on Jan 1st 2013 = 50c)
- Potential information market conditions (current stock price)
 - ☐ 40c -> Stock undervalued -> Buy
 - ☐ 60c -> Stock overvalued -> Sell
 - ☐ 50c -> Meets expectation -> Hold

Figure 9.8.: Experiment 1 - Instructions page 2

First, an estimation target for the market is defined (This example question is made up):

“What will be the market share of the Apple iPhone 5 on January 1st 2013?”

Before the market starts, we will ask you for an initial estimation of the true value just like in the first part of the experiment.

Second, the estimation target is transformed into an information market stock:

“Market share of the Apple iPhone 5 on January 1st 2013”

Third, the information market is started.

When the market starts, all participants are given the same number of stocks (1000 in our markets) and an amount of information market currency (1000 e-Cents in our experiment) so they can buy and sell stocks right from the start. All traders enter the information market at the same time. The real underlying value will be somewhere between 0 and 100 percent. Consequently, the stock price may be traded anywhere between 0 and 100 e-Cents.

- When a trader decides to buy a stock, the stock price will automatically move upwards.
- When a trader sells a stock, the stock price will automatically be adjusted downwards.
- Traders can trade in the market for 140 seconds.

Fourth, the information market is closed

After the market is closed, all stocks will be paid out corresponding to the real underlying percentage value. That means: If you think, the stock price is too low you should consider buying stock. If you think, the stock price is too high you should consider selling stocks. If you are not sure or uncertain, you can observe, how other traders act in the market when the stock price changes, learn from their actions and consider acting accordingly.

You will be asked to give a revised estimate of the true value after the market has closed. You will be trading in a total of 6 information markets, each one with a different underlying estimation task in the same product category and market. Each of the markets is totally independent from the other markets. So whenever a new market starts, you will start out with a new set of 1000 stocks and a new cash balance of e-Cents, no matter how you performed in the previous markets. You will not be shown the true value of each prediction target before all six markets have been finished.

Figure 9.9.: Experiment 1 - Instructions page 3

Pre-experimental questionnaire

Your experiment number: _____

	<i>Fully agree</i>				<i>Fully disagree</i>
I trade or plan to trade shares, bonds or derivatives actively in the stock market.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I am very interested in what is happening in the stock markets.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<i>Fully agree</i>				<i>Fully disagree</i>
I am very good at estimating the outcome of events (sports, bets) even if I do not possess enough information to know the real outcome.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Compared to the people around me, I am usually better at estimating unknown or future outcomes of events.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<i>Fully agree</i>				<i>Fully disagree</i>
Among all kinds of advertisements, consumer electronics advertisements are particularly interesting to me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Given my personal product interests, the consumer electronics market is very relevant to me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Overall, I am quite involved when I am purchasing new consumer electronics for my own personal use.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
How likely would the following be	<i>Very likely</i>				<i>Very unlikely</i>
Investing 10% of your annual income in a moderate growth mutual fund (Aktienfond).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Investing 5% of your annual income in a very speculative stock.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Investing 10% of your annual income in a new business venture.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 9.10.: Experiment 1 - Pre-experimental questionnaire

Evaluation Questions for Experimental Manipulation

1. How successful is the Mini compared to VW Polo and Fiat Punto? If the sales of all three models make up for 100%, what would be the Mini's percentage market share in the first quarter of 2011?
2. By how many percent did the average CO2 Emission per kilometer regress among newly registered cars from 2006 to 2011 in Germany?
3. The class of compact cars (like VW Golf, Ford Focus) is generating the most car revenues in Germany in the second quarter 2011. How many percent of all newly registered cars belong to the compact class?
4. At least in media, cars with electric drives are prominently featured. By what percentage did the amount of registered electric vehicles increase from 1.1.2010 to 1.1.2011 in Germany?
5. The top selling Volkswagen models in May 2011 are Polo, Passat, Golf and Tiguan. How many percent of new registrations of these four vehicles are accounted for by the VW Passat?
6. On January 1st 2011, what percentage of cars in Germany was registered to women?
7. German cars are also manufactured for international markets. Of all cars manufactured in Germany in 2010, how many percent were exported to other European countries?
8. How many percent of cars were registered to private individuals (and not to business entities) in Germany on 1.1.2011?
9. By how many percent did the Volkswagen Passat registrations in Germany increase in July 2011 compared to the same month in 2010?
10. 86811 cars by the two big Italian manufacturers Fiat and Alfa Romeo were registered in 2010. What percentage of these registrations is accounted for by the more expensive Alfa Romeo brand?

Evaluation Questions for Experimental Information Markets

1. In the first quarter of 2011, 1.026.000 cars were newly registered in Germany. What percentage of these cars was made by German manufacturers?
2. The Golf Plus is a novel and roomier version of Germany's most famous car: The Volkswagen Golf. 30-40 year olds are even referred to as the generation "Golf", as they grew up with the car and it is very popular with them. What percentage of the new Volkswagen Golf Plus were actually bought by individuals older than 60 in the first quarter of 2011 in Germany?
3. Over the last years, an increasing number of the newly registered cars were painted white. What percentage of all newly registered cars in Germany in 2010 was really painted white?
4. The new Mercedes SLK was introduced in 2011. By what percentage did the sales figures for the Mercedes SLK increase in the first Quarter in 2011 compared to the first quarter in 2010?
5. How many percent of the people who so far bought the new Fiat 500 are female?
6. Prediction market question: By what percentage did the amount of newly registered hybrid cars like the Toyota Prius increase from 1.1.2010 to 1.1.2011 in Germany?

Figure 9.11.: Experiment 1 - Evaluation questions and information-market questions

After-experimental questionnaire

Your experiment number: _____

	<i>Fully agree</i>				<i>Fully disagree</i>
I enjoyed participating in the experiment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Participating in the experiment was fun.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I understood the tasks in the experiment	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	<i>Fully agree</i>				<i>Fully disagree</i>
I estimated better than the other participants in the experiment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
My profits in the information market are higher than the other participants' profits.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Compared to the other participants of the experiment, how high do you think your total profit is?
 (0 = my profit is among the lowest 5%; 100 = my profit is among the top 5%) _____

My performance in the experiment can be mainly attributed to...	<i>Fully agree</i>				<i>Fully disagree</i>
...my capability to estimate unknown values.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...my knowledge in the product domain.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... (bad) luck with the specific questions in this experiment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	<i>Native</i>	<i>Intermediate</i>	<i>Beginner</i>
My English skills	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 9.12.: Experiment 1 - Post-experimental questionnaire page 1

Market understanding (please circle the appropriate reply)

1. If the stock price is currently 80 e-Cents and I strongly think that the real value is below 40 percent, what should I do?
(Buy shares) (Observe the market) (Sell shares)
2. If the stock price is currently 20 e-Cents and I strongly believe that the real value is above 10 percent but below 20 percent, what should I do?
(Buy shares) (Observe the market) (Sell shares)
3. If the stock price has just increased from 30 e-Cents to 45 e-Cents and I believe that the real value is between 40 and 50 e-Cents, what should I do?
(Buy shares) (Observe the market) (Sell shares)

My age: _____

My gender: male / female

My place of origin: _____

My e-mail for a chance to win one of the vouchers: _____

Instructions for the Experiment

Introduction

This experiment is concerned with a mechanism that can be used to evaluate innovative products: information markets. The experiment will consist of two parts. Both parts offer you the chance to win an Amazon voucher. **2* 20,- Euros in the first part** of the experiment and **12 * 20,- Euros in the second part** of the experiment.

In the first part, we are going to test your estimation performance in a series of domain specific estimation tasks. Afterwards you will be asked for your personal self-assessment.

In the second part, you will compete with other individuals in an information market, where you can generate profits from your estimations in a market environment.

Before the Live – Experiment, you will get to view a detailed instruction video.

Basic and important rules

1. Try to estimate, guess or trade as well as you can. If you deliberately enter false and implausible values, you are lowering your chance of winning and you are making the scientific results useless.
2. If you are lacking information regarding the questions, use your gut feeling.
3. During the experiment, please stay at your booth, leave your headphones on and do not cheat. You would deprive yourself of the chance to win and make the scientific results useless for us.
4. After the experiment, please do not talk to other and later participants about the experiment, the tasks and your answers and performance. Again, you would deprive yourself of the chance to win and make the scientific results useless for us.
5. Thank you for your participation and have fun!

Part 1: Estimation task

- You will answer a set of 10 estimation questions
- The questions will deal with a specific market of well-known products
- After the 10 questions, you will be asked to state how many questions you believe to have answered correctly.
- Among all subjects, we will raffle a **20,-€Amazon** voucher. Your probability to win the voucher will highly depend on two things: 1. Answer as many questions as possible correctly, 2. Try to estimate your own performance as accurately as possible.

Part 2: Information market

In an information market, outcomes of uncertain events are traded by the participants. In our case, the outcomes of questions regarding innovations and new products in a well-known market will be traded.

First, an estimation target for the market is defined (This example question is made up):

“What will be the market share of the Apple iPhone 5 on January 1st 2013?”

Before the market starts, we will ask you for an initial estimation of the true value just like in the first part of the experiment.

Second, the estimation target is transformed into an information market stock:

“Market share of the Apple iPhone 5 on January 1st 2013”

Third, the information market is started.

When the market starts, all participants are given the same number of stocks (1000 in our markets) and an amount of information market currency (1000 e-Cents in our experiment) so they can buy and sell stocks right from the start. All traders enter the information market at the same time. The real underlying value will be somewhere between 0 and 100 percent. Consequently, the stock price may be traded anywhere between 0 and 100 e-Cents.

- When a trader decides to buy a stock, the stock price will automatically move upwards.
- When a trader sells a stock, the stock price will automatically be adjusted downwards.
- Traders can trade in the market for 140 seconds.

Fourth, the information market is closed

After the market is closed, all stocks will be paid out corresponding to the real underlying percentage value. That means: If you think, the stock price is too low you should consider buying stock. If you think, the stock price is too high you should consider selling stocks. If you are not sure or uncertain, you can observe, how other traders act in the market when the stock price changes, learn from their actions and consider acting accordingly.

You will be asked to give a revised estimate of the true value after the market has closed. You will be trading in a total of 6 information markets, each one with a different underlying estimation task in the same product category and market. Each of the markets is totally independent from the other markets. So whenever a new market starts, you will start out with a new set of 1000 stocks and a new cash balance of e-Cents, no matter how you performed in the previous markets. You will not be shown the true value of each prediction target before all six markets have been finished.

The Information Market Interface:

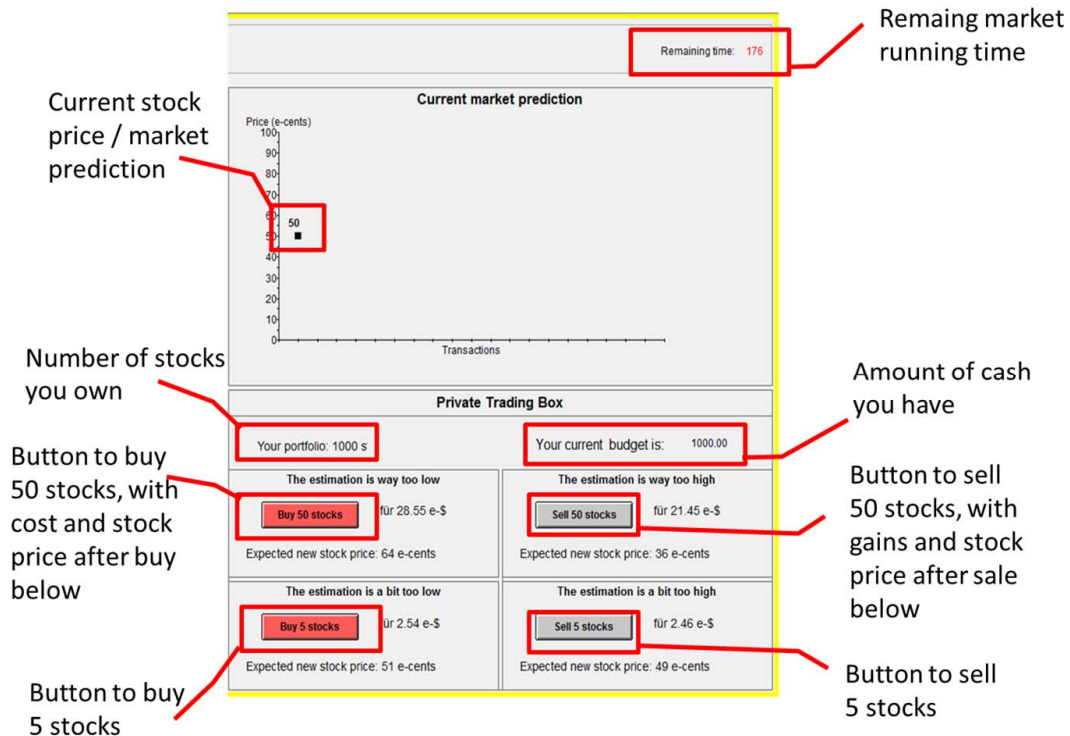
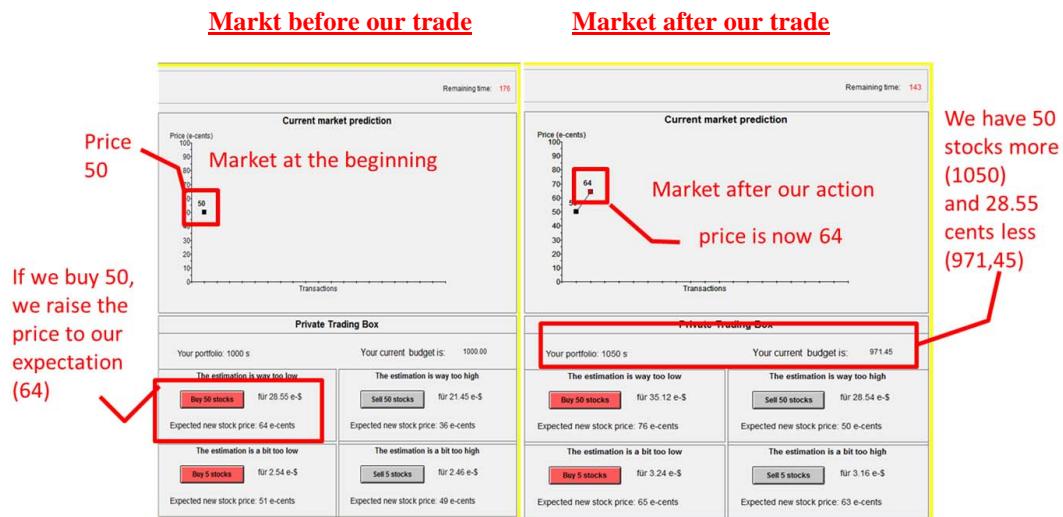


Figure 9.16.: Experiment 2 - Instructions page 3

Information Market Example 1: We think the real value will be higher than the current stock price / market expectation

- Let's assume the prediction task is still „What will be the market share of the iPhone 5 in January 2013? “
- Now let's assume we believe, the market share will be 64%. What should we do in the information market?
- Based on our belief, every stock will be worth 64 cents after the market. At the beginning (left screenshot) they cost 50 cents. That means we should buy stocks, if we are convinced enough that the real market share will be 64%.
- We pay roughly $50 \text{ stocks} * (64+50)/2 \text{ e-cents} = 28.5 \text{ e-cents}$
(There is a complicated function behind the mechanism, which should not concern you)



- Now let's assume the market finishes and the true market share of the iPhone in January 2013 turns out to be 65% (almost like we believed). How did we profit from buying the 50 stocks?

Our market profit **before** buying 50 stock: $1000 \text{ stocks} * 0,65 + 1000 = 1650$

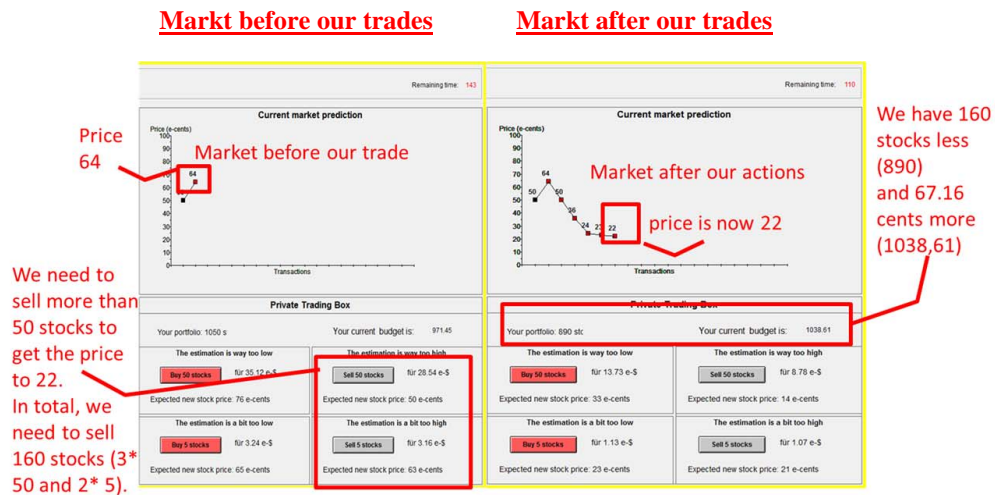
Our market profit **after** buying 50 stock: $1050 \text{ stocks} * 0,65 + 971,45 = 1653,95$

We made a **profit** by trading in the market and at the same time, we made the market prediction / stock price more **accurate**.

Figure 9.17.: Experiment 2 - Instructions page 4

Information Market Example 2: We think the real value will be lower than the current stock price / market expectation

- Let's assume the prediction task is still „What is the market share of the iPhone 5 in January 2013? “
- Now let's assume we changed our expectation because new information came to our mind (We remembered all the great Samsung mobiles coming to the market). We now believe the market share of the new iPhone5 will only be 22%. What should we do in the information market?
- Based on our belief, every stock will be worth 22 cents after the market. Right now (left screenshot) they cost 64 cents. That means we should sell stocks, if we are convinced enough that the real market share will be 22%.



- Now let's assume the market finishes and the true market share of the iPhone in January 2013 turns out to be 25% (we were right about the impact of new Samsung models). How did we profit from selling the 160 stocks?

Our market profit **before** selling 160 stock: $1050 \text{ stocks} * 0,25 + 971,45 = 1233,95$

Our market profit **after** selling 160 stock: $890 \text{ stocks} * 0,25 + 1038,61 = 1261,11$

We made a **profit** by trading in the market and at the same time, we made the market prediction / stock price more **accurate**.

Figure 9.18.: Experiment 2 - Instructions page 5

Central take-away lessons for the experiment:

1. If you think the current stock price / market prediction is **lower than the real value** you should **buy stocks**.
2. If you think the stock price / market prediction is **higher than the real value**, you should **sell stocks**.
3. If you are **uncertain about the real value**, you should **observe the market price and only buy or sell, when you reach sufficient certainty**.

Enjoy the experiment!

Pre-experimental questionnaire

Your experiment number: _____

	<i>Fully agree</i>				<i>Fully disagree</i>
I trade or plan to trade shares, bonds or derivatives actively in the stock market	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I am very interested in what is happening in the stock markets	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<i>Fully agree</i>				<i>Fully disagree</i>
I am very good at estimating the outcome of events (sports, bets) even if I do not possess enough information to know the real outcome.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Compared to the people around me, I am usually better at estimating unknown or future outcomes of events.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<i>Fully agree</i>				<i>Fully disagree</i>
Among all kinds of advertisements, consumer electronics advertisements are particularly interesting to me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Given my personal product interests, the consumer electronics market is very relevant to me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Overall, I am quite involved when I am purchasing new consumer electronics for my own personal use.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
How likely would the following be	<i>Very likely</i>				<i>Very unlikely</i>
Investing 10% of your annual income in a moderate growth mutual fund (Aktienfond).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Investing 5% of your annual income in a very speculative stock.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Investing 10% of your annual income in a new business venture.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 9.20.: Experiment 2 - Pre-experimental questionnaire

Evaluation Questions for Experimental Manipulation

1. What was the market share of the Macbook Air among ultra-thin notebooks in the first half of 2012 in Germany?
2. What was the market share of the Windows Mobile operating system among all smartphone operating systems in June 2012 in Germany?
3. By what percentage did the battery time of ultra-thin notebooks increase from June 2012 to November 2012 in Germany?
4. What was the global market share of Google's Chrome web browser in the third quarter 2012?
5. How many percent of all laptops sold in the first half of 2012 in Germany had an SSD (Solid-state disk) drive included instead of a traditional hard drive?
6. How many percent more cumulative global downloads did the Apple App store have compared to the Google Play store by September 2012?
7. What is the market share of Nokia in total in the emerging mobile hand-set markets of the African continent in the first half of 2012?
8. How many percent of German mobile phone owners used smartphones in the first half of 2012?
9. How many percent of the US internet users owned tablet pcs by the first half of 2012?
10. Samsung and LG are Korean LCD-Screen manufacturers that dominate global screen markets. What was their combined market share in the global LCD flat-panel market in the third quarter 2012?

Evaluation Questions for Experimental Information Markets

1. What will be the market share of 7 inch tablets among all media tablet pcs in Germany in January 2013?
 - a. Info 1: Among all Samsung tablets sold between January and June 2012, the market share of 7 inch tablets was 13% and slightly higher than the industry average.
 - b. Info 2: In June, 7 inch tablets had a market share of 8.5% in Germany.
 - c. Info 3: Between June and November 2012, the market share of 7 inch tablets more than tripled in Germany.
2. How many percent of all tablets sold in January 2013 in Germany will be made by Apple?
 - a. Info 1: In June 2012 in Germany, seven of the most commonly sold tablets were made by Apple
 - b. Info 2: Apple's market share among tablets dropped by almost 40% between June 2012 and November 2012 in Germany.
 - c. Info 3: In October 2012, Apple's tablet market share dropped below 50% in Germany, but Apple was still strong market leader.

Figure 9.21.: Experiment 2 - Evaluation questions and information-market questions

3. How many percent of all tablets sold in January 2013 in Germany will have 3G / 4G broadband internet access?
 - a. Info 1: Among the more expensive tablets sold between 400 and 700 Euro in October 2012 in Germany, 61.7% had a 3G/4G module.
 - b. Info 2: Among the cheaper tablets sold between 0 and 300 Euro in October 2012, 10.7% had a 3G/4G module.
 - c. Info 3: In May 2012, the market share of tablets with 3G/4G modules was 60% but has continually decreased since then. More modules with WIFI only such as the Amazon Kindle Fire or the Google Nexus captured much market share.
4. What will be the share of Android among Smartphone operating systems in January 2013 in Germany?
 - a. Info 1: The iPhone with the IOS operating system is by far the best-selling single phone in Germany in any month in 2012.
 - b. Info 2: Android's market share is below its average in low price categories. For units below 200€ sales price, Android's market share was 61.4% in October 2012.
 - c. Info 3: The android market share peaked at 78.3% in August 2012 but dropped until October by more than 10% due to the new iPhone introduction.
5. By how many percent will the prices for 128GB SSD hard drives (Solid-state drives) drop in Germany between January 2012 and January 2013?
 - a. Info 1: Traditional hard-drives sold for an average of 119 Euro in January 2012 and 102 Euro in November 2012, which means a 15% price decrease. Traditional hard-drive prices decreased less than SSD storage prices.
 - b. Info 2: The average Price for a 128GB SSD was 183 Euros in January 2012 in Germany.
 - c. Info 3: By October 2012, average 128GB SSDs prices were below the average prices for traditional hard drives in Germany.
6. How many percent of all newly sold TFT-screens in Germany will have a screen resolution bigger than 22 inch in January 2013 in Germany?
 - a. Info 1: 24 inch monitors had a market share of 22% in October 2012 in Germany
 - b. Info 2: In January 2011 in Germany, the market share of LCD-Monitors bigger than 22 inch was 35.8%.
 - c. Info 3: In January 2012 in Germany, the market share of LCD-Monitors bigger than 22 inch was 44.3%.

Figure 9.22.: Experiment 2 - Evaluation questions and information-market questions continued

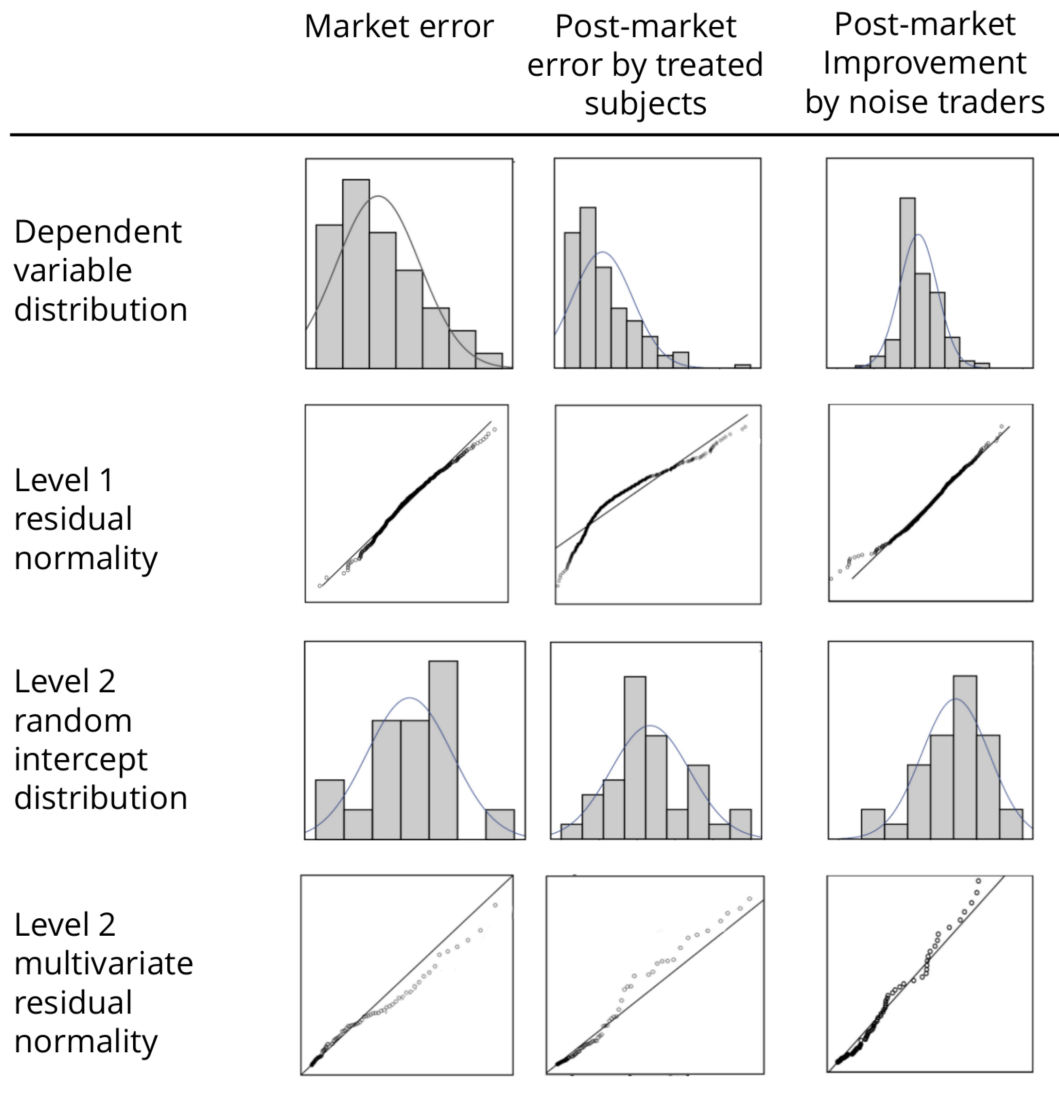


Figure 9.23.: Experiment 2 - Data adequacy check for the basic experimental setting

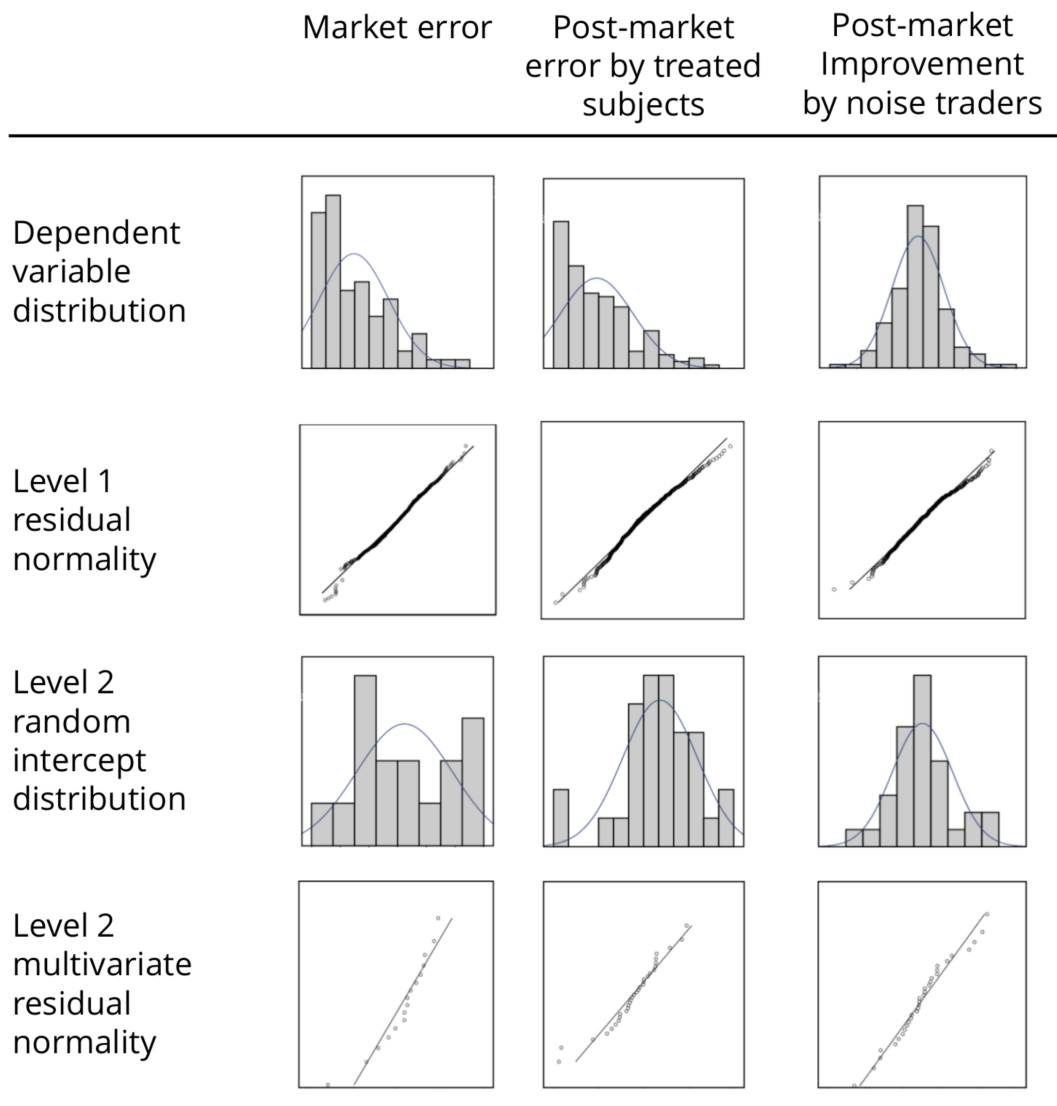


Figure 9.24.: Experiment 2 - Data adequacy check for the additional experimental setting

After-experimental questionnaire

Your experiment number: _____

	<i>Fully agree</i>				<i>Fully disagree</i>
I enjoyed participating in the experiment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Participating in the experiment was fun.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I found the experiment more enjoyable than the other tasks during the B2B-Marketing lecture.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Compared to the other participants of the experiment, how high do you think your total profit is?
(0 = my profit is among the lowest 5%; 100 = my profit is among the top 5%) _____

	<i>Fully agree</i>				<i>Fully disagree</i>
I estimated better than the other participants in the experiment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
My profits in the information market are higher than the other participants' profits.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	<i>Fully agree</i>				<i>Fully disagree</i>
Predicting the market share of consumer electronic products in Germany is very hard for me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I would have to study very hard to predict market shares of consumer electronics products in Germany well .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Compared to most other students in my class, predicting the market share of consumer electronic products in Germany is very hard for me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I would have to study very hard to predict consumer electronics products in Germany better than other students in my class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 9.25.: Experiment 2 - Post-experimental questionnaire page 1

After-experimental questionnaire

Your experiment number: _____

	<i>Fully agree</i>				<i>Fully disagree</i>
I enjoyed participating in the experiment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Participating in the experiment was fun.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I found the experiment more enjoyable than the other tasks during the B2B-Marketing lecture.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Compared to the other participants of the experiment, how high do you think your total profit is?
(0 = my profit is among the lowest 5%; 100 = my profit is among the top 5%) _____

	<i>Fully agree</i>				<i>Fully disagree</i>
I estimated better than the other participants in the experiment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
My profits in the information market are higher than the other participants' profits.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	<i>Fully agree</i>				<i>Fully disagree</i>
Predicting the market share of consumer electronic products in Germany is very hard for me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I would have to study very hard to predict market shares of consumer electronics products in Germany well .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Compared to most other students in my class, predicting the market share of consumer electronic products in Germany is very hard for me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I would have to study very hard to predict consumer electronics products in Germany better than other students in my class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 9.26.: Experiment 2 - Post-experimental questionnaire page 2

Part VII.

Curriculum Vitae

Jan-Paul Lüdtkke

Contact information

Holstenstraße 110

22767 Hamburg

Phone: 0049-176-20112541

Email: jp.luedtke@gmail.com

Personal details

Born: June 5, 1981 in Bremen, Germany

Nationality: German

Current position

Founder and Managing Director, Akanoo GmbH, Hamburg

Work experience

2008-2009 – Marketing Consultant, GfK SE, Nürnberg

2009-2013 – Research Assistant, Hamburg University of Technology, Hamburg

2013-today – Startup Manager, Thomas JC Matzen GmbH, Hamburg

2013-today – Founder and Managing Director, Akanoo GmbH, Hamburg

Education

2008 – DIPLOMA in Business Administration, Universität Passau