Psychological Process Models

and Aggregate Behavior

Dissertation

zur Erlangung des Akademisches Grades

doctor rerum naturalium (Dr. rer. nat.)

im Fach Psychologie

eingerichet an der

Mathematisch-Naturwissenschaftlichen Fakultät II

der Humboldt-Universität zu Berlin

von

Pantelis Pipergias Analytis, M.Sc.

Prasident der Humboldt-Universität zu Berlin

Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissentschaftlichen Fakultät II

Prof. Dr. Elmar Kulke

Gutacther:

1.Prof. Dr. Gerd Gigerenzer

2.Prof. Dr. Matthias Ziegler

3.Prof. Dr. Gaël Le Mens

Tag der Einreichung: 03. December 2014

Tag der Verteidigung: 15. Juni 2015

Contents

Summary **vi**

Zusammenfassung **viii**

1 Introduction **1**

1.1 Stories of convergence 2

1.1.1 Behavioral economics 3

1.1.2 Computational social science 4

1.2 Overview of models presented in this dissertation 5

1.3 References 8

2 Multi-attribute utility models as cognitive search engines **11**

2.1 Introduction 12

2.2 The theoretical framework 15

2.2.1 The environment 15

2.2.2 The optimal strategy 15

2.2.3 Subjective models 16

2.2.4 An example 17

2.3 Results 19

2.3.1 Real world environments 20

2.3.2 Performance 21

2.3.3 Impact of search order on performance and search length 24
2.3.4 The role of the standard deviation σ 27
2.3.5 Paired-comparison results 30
2.4 Discussion 32
2.4.1 Conceptual implications of our framework 32
2.4.2 Applied implications of our framework 33
2.4.3 Connections to prescriptive and descriptive decision making 35
2.4.4 Extensions and limitations 37
2.4.5 Conclusion 39
2.5 References 40
2.6 Appendix 1: Proof of result 1 47

3 Social influence and the collective dynamics of opinion formation 48
3.1 Introduction 50
3.2 Experiments 53
3.2.1 Experimental design 53
3.2.2 Experimental results 54
3.3 Simulations 60
3.3.1 The model 60
3.3.2 Collective dynamics results 62
3.4 Discussion 66
3.5 References 70
3.6 Appendix: Supplementary material 74
3.7 Appendix 2: Materials 76

4 The collective dynamics of sequential search in markets for cultural products 81

4.1 Introduction 83
4.1.1 Literature review 85
4.1.2 Our modeling approach 89
4.2 The model 92
4.2.1 The environment 92
4.2.2 Random search 93
4.2.3 Social interaction 94
4.2.4 Popularity heuristic 95
4.2.5 Assumptions 96
4.3 Results 99
4.3.1 Utility, popularity and inequality 100
4.3.2 Outcome unpredictability and luck 103
4.3.3 Welfare analysis 105
4.4 Discussion 110
4.4.1 Luck and dynamics of success in cultural markets 111
4.4.2 Exploration, imitation and diversity 112
4.4.3 Revisiting the Music lab experiment 114
4.4.4 Managing social influence 115
4.4.5 Improving the external validity 115
4.4.6 Concluding remarks 116
4.5 References 118

5 General discussion **125**
5.1 Integrating cognitive and agent-based models 126
5.2 Psychological process models and prediction 127
5.3 The value of simplicity and applicability 129

5.4 Correspondence rather than coherence 129

5.5 Concluding remarks 130

5.6 References 131

Summary

This dissertation comprises of three independent essays which introduce novel psychologically inspired process models and examine their implications for individual, collective or market behavior.

The first essay studies multi-attribute choice as a guided process of search. It puts forward a theoretical framework which integrates work on search and stopping with partial information from economics with psychological subjective utility models from the field of judgment and decision making. The alternatives are searched in order of decreasing estimated utility, until the expected cost of search exceeds the relevant benefits; The essay presents the results of a performance comparison of three well-studied multi-attribute choice models, (i) the multi-attribute linear utility, (ii) the equal weighting of alternatives and (iii) a single attribute heuristic on twelve real world datasets. It explains the role of the preference ranking and the estimation error of the models in their performance across real-world environments.

The second essay reports the results of two experiments designed to understand how people revise their judgments of factual questions after being exposed to the opinion and confidence levels of others. It introduces a tree model of judgment revision which is directly derived from the empirical observations. The model demonstrates how opinions in a group of interacting people can converge or polarize over repeated interactions. Two major attractors of opinion are identified (i) the expert effect induced by the presence of a highly confident individual in the group and (ii) the majority effect caused by a critical mass of people sharing similar opinions. Additional simulations show the existence of a tipping point at which one attractor will dominate the other, driving collective opinion in a given direction.

The third essay, studies collective behavior in markets for search products. It is assumed that decision makers with diverse yet correlated preferences decide sequentially which among numerous alternatives to choose. The decision makers consider the alternatives in order of decreasing popularity and choose the first alternative with utility higher than a certain satisficing threshold. The popularity order is updated after each individual choice. The presented framework illustrates that such markets are characterized by rich-get-richer dynamics which lead to inequality in the market-share distribution and unpredictability in regard to the final outcome. The satisficing threshold employed by the decision makers and the preference diversity in the decision maker population determine the strength of these effects.

Keywords: search, process models, heuristics, social influence.

Zusammenfassung

Diese Dissertation umfasst drei voneinander unabhängige Artikel. In diesen werden neue Prozessmodelle vorgestellt, die von der entscheidungspsychologischen Forschung inspiriert wurden. Sie liefern weiterführende Erkenntnisse zu individuellem Verhalten, Märkten und kollektivem Verhalten.

Im ersten Artikel werden Entscheidungsprozesse mit mehreren Entscheidungsmerkmalen (sog. multi-attribute choices) als gesteuerte Suchprozesse modelliert. Zunächst wird ein theoretischer Rahmen vorgestellt, in dem ökonomische Modelle Entscheidungen mit Suche mit Modellen des subjektiven Nutzens aus dem Bereich der psychologischen Forschung zum Urteilen und Entscheiden integriert wird. In den so modellierten Entscheidungsprozessen wird angenommen, dass Individuen ihre Entscheidungsalternativen nach deren abnehmenden Nutzen ordnen und dann so lange durchsuchen, bis die erwarteten Suchkosten höher als die entsprechenden Gewinne sind. Anschlieend wird die Güte dreier multi attribute - Entscheidungsmodelle an zwölf realen Datensätzen überprüft. Diese Modelle sind (i) ein multi-attribute-Modell, dass einen linear absteigenden Nutzen der attributes annimmt (ii) ein Modell, welches gleichen Nutzen für alle attributes annimmt und schlielich (iii) eine Heuristik, die allein ein attribute verwendet. Die Rolle von Präferenzreihenfolgen sowie die des Schätzfehlers aller drei Modelle werden anschlieend diskutiert.

Im zweiten Artikel werden die Ergebnisse zweier Experimente vorgestellt, in denen untersucht wurde, wie Personen ihre Urteile verändern, wenn sie den Urteilen und dem der Konfidenzniveau anderer Personen ausgesetzt sind. Ein Baummodell wird eingeführt, welches abbildet, wie Urteile aufgrund solcher Informationen revidiert werden. Dieses Modell basiert auf den Ergebnissen der beiden Experimente: Indem soziale Informationen berücksichtigt werden, kann es zeigen, wie Urteile in einer Gruppe interagierender Personen zusammenlaufen bzw. polarisieren, wenn diese wiederholt miteinander interagieren. Es werden zwei Faktoren identifiziert, die Urteile Einzelner beeinflussen: i) die Rolle von Experten sehr konfidenter Personen in einer Gruppe und ii) die Rolle von Mehrheiten, das heißt, einer kritischen Anzahl anderer, die jeweils die gleiche Meinung haben. In zusätzlichen Simulationen wird gezeigt, an welchen Kipppunkten ein Faktor den anderen jeweils dominiert und dann die kollektiven Urteile in eine bestimmte Richtung lenkt.

Im dritten Artikel wird kollektives Verhalten in Märkten für kulturelle Produkte untersucht. Es wird angenommen, dass Personen mit unterschiedlichen, aber miteinander korrelierten, Präferenzen sequentiell entscheiden, welche von vielen Optionen sie wählen. Personen ordnen die Optionen entsprechend ihrer Popularität an und entscheiden sich dann für diejenige, die einen Nutzen hat, der über einer bestimmten ausreichend guten Schwelle liegt. Nach jeder individuellen Entscheidung wird die Rangfolge revidiert. Innerhalb dieses einfachen Rahmens wird demonstriert, dass solche Märkte durch eine sogenannte rich get richer-Dynamik charakterisiert sind. Diese führt zu Ungleichheiten in den Marktanteilen und ungewissen finanziellen Erlösen. Die Stärke solcher Effekte wird durch die Vielfalt der Präferenzen auf dem Markt sowie den Nutzen-Schwellenwert einzelner Personen bestimmt.

Schlagwrter: Suche, Prozessmodelle, Faustregeln, sozialer Einfluss

List of Figures

2.1 An example 18

2.2 Model performance 23

2.3 Returns to search 26

2.4 Model search length 29

3.1 Initial distribution of opinions - relation between confidence and opinion accuracy 55

3.2 Wisdom of the crowds 56

3.3 Influence map 58

3.4 Confidence revision 59

3.5 Collective dynamics 63

3.6 Majority vs experts 65

3.7 Distribution of answers for all 32 questions 74

3.8 Confidence dynamics 75

3.9 Screenshot of the first experiment 79

3.10 Screenshot of the second experiment 80

4.1 Utility and popularity 101

4.2 Market share inequality 103

4.3 Unpredictability 104

4.4 Welfare analysis 106

4.5 Welfare dynamics 108

List of Tables

2.1 The environments 20

2.2 Average model performance 21

2.3 Average model search 27

2.4 Paired-comparison results 31

3.1 Questions and answers 77

Acknowledgements

The present dissertation is the outcome of four years of work as a fellow of the International Max Planck Research School on Adapting Behavior in a Fundamentally Uncertain World (Uncertainty School). During this time I was based at the Adaptive Behavior and Cognition (ABC) Group at the Max Planck Institute for Human Development in Berlin.

The interdisciplinary nature of the ABC Group, and the Uncertainty School more broadly, brought me in contact with a diverse pool of ideas and a culture of critical thinking and debate. This allowed me to quickly develop my scientific personality. I would like to thank Gerd Gigerenzer, the director of the ABC group, for giving me the opportunity to be part of the Max Planck Institute family.

As a member of the ABC Group I was closely supervised by Konstantinos Katsikopoulos, who was always there with advice and support. The outcome of our long discussions was finally consolidated in the project presented in the second chapter of this dissertation. Besides being my supervisor, Konstantinos became a very good friend. Thank you Konstantinos for all your support.

The research life at the ABC Group was filled with the stimulating exchange of ideas with many of my colleagues. In the course of the first and second year of my Phd I teamed up with Juliane Kammer, Mehdi Moussaid, and Hansjorg Neth and studied social influence and collective behavior. Some of the results of our research are reported in chapter two of the current dissertation. Juliane, Mehdi and Hans, I am grateful for the opportunity to work with each one of you.

I was fortunate to work a few steps away from Amit Kothiyal, with whom I share many interests. It was only a matter of time before we collaborated on a scientific project. Out of the many things I learned from Amit, Emacs is foremost, in which I now spend most of my computer time.

I am thankful to the co-ordinators and professors of the Uncertainty School, and in particular to Susanne Büchner for the summer schools and workshops that she organized and to Werner Güth and Oliver Kirchkamp for their scientific advice. Over the course of one of the summer schools I first met and worked together with Hrvoje Stojic. I now collaborate with Hrvoje regularly and share many common research plans. Hrvoje, thanks for your feedback and your advice.

Everyday life at the institute has been filled with great people and lots of laughter. Perke, Malte, Jon, Florian, Nadine, Wasi, Stojan, Astrid, Nicolai, Azzura, Jan, Michelle, Cristian and everybody else in the ABC group thank you for all the great moments.

A big thanks to my family in Greece for their patience and support. This thesis is dedicated to you and in particularly to my brothers Aggelos and Iasonas who will be soon starting their University studies. Finally, thanks to everybody in Thermofix and Filmarche for making Berlin feel like home.

Chapter 1

Introduction

Although psychology is the science of behavior par excellence, it has had so far only a modest contribution to the study of aggregate behavior. In the past, neighboring disciplines such as economics, sociology, and more recently physics, have advanced the conceptualization of aggregate behavior, yet in most cases with little regard to the psychological plausibility of the models advanced to describe individual agents. Herbert Simon (1998) wittingly exposes this paradox in an essay on the relation of psychology and economics. Simon points out that if psychology is the science of human behavior, and economics the science that deals with behavior in the marketplace, then economics appears to be a branch of psychology. The same remark can be made for other social sciences such as sociology and political science. The phenomena that they study are the product of human behavior. However, there are only limited cases in which human psychological processes are directly invoked to explain or predict aggregate level observations. In fact, today most models of economic and social phenomena are still based on descriptions of individual agents that do not correspond to the state-of-the-art models studied by psychologists.

Instead of psychological theories, the most expansive model of individual behavior in the social sciences in the second half of the 20th century has been rational choice theory, which in most of its expressions postulates that decision makers are utility maximizers with herculean computational capacities. The development of rational choice theory has been primarily driven by the internal coherence of its assumptions rather than by correspondance to actual behavior observed in real-world environments. Thus, other than the motivational element, expressed in the concept of utility, the theory as practiced by social scientists, has little psychological content. Besides, there is a striking schism between theoretical rational choice models and empirical practice. The original theories are in many cases replaced by naive statistical models, which are used to explain or to predict aggregate behavioral measures. Yet these models do not describe any actual processes and have little structural connections to the theory itself (Papandreou, 1958; Shmueli, 2010).

Forming accurate beliefs about the economy and the social world is vital for both individuals and organizations. Thus, it is not surprising that social scientists regularly advise governments and firms. Although rational choice theory provides a structured modeling framework, in most cases models of this type have limited correspondence to actual real world problems. Further, while it is true that statistical models agnostic to the exact cognitive or social processes can be directly applied to derive reasonable predictions about individual or aggregate behavior, such models do not provide true insights into the processes that link micro to macro-behaviors. Understanding these processes is a key factor in order to intervene efficiently in a system and improve its possible outcomes. By capitalizing on psychological methods and theories, we can design models which are based on sound micro-foundations at the level of the individual agents and explicitly take into account the processes that lead from individual to aggregate behavior.

1.1 Stories of convergence

Over the last decades two new fields of research, behavioral economics and computational social science, have developed independently of each other and have opened the way for a new conceptualization of the relation between individual and aggregate level behavior. Behavioral economics has departed from rational choice theory seeking psychological foundations for economic behavior. Computational social science, in contrast, has its origins in agent-based models with minimal assumptions about individual behavior (Schelling, 1978). Today, it offers a host of new methodological possibilities for integrating psychology with the social sciences.

1.1.1 Behavioral economics

Already in the 1950s, Ward Edwards (1954) and his associates started to investigate experimentally the cognitive processes related to decisions in choice contexts such as those described in economics. Across several instances, Edwards found that actual human behavior did not correspond to the assumptions put forward by rational choice theory (e.g. Phillips and Edwards, 1966). At the same time, Herbert Simon (1955; 1956) sought an alternative more realistic conceptual framework to define human rationality. He advanced the view that people are boundedly rational and their cognitive strategies are adapted to the structure of the environment within which they act. Simon's approach has been further pursued in organization science (Cyert and March, 1963) and psychology (Gigerenzer and Todd, 1999), yet as a heterodox alternative alongside the mainstream rational choice theory. The work of Edwards, Simon, and their collaborators has also paved the way for the development of behavioral economics.

The beginning of behavioral economics is mostly associated with the work of the psychologists Amos Tversky, who was a student of Edwards, and Daniel Kahenmann in the early 70s (Heukelom, 2009). Kanhmann and Tversky systematically exposed cases where decision makers violated the norms of rational choice theory. In addition, they postulated psychological alternatives to economic models. For example, they advanced prospect theory as an alternative to expected utility theory to describe how people make decisions in choice problems under risk (Kahneman and Tversky, 1979). At approximately the same time, economists added experiments to their methodological repertoire. By now there have been several successful cases where psychological theories have been applied directly to theoretical scenarios developed within the rational choice framework. Erev and Roth (1998), as an example, have successfully used learning models developed by psychologists (e.g. Estes, 1950) to predict choices in strategic games.

In another line of research, psychological models and findings have been employed to directly explain behavior at the aggregate level in groups, markets, or the economy as a whole. Benartzi and Thaler (1995), for example, used a myopic version of cumulative prospect theory (Tversky and Kahneman, 1992) to explain the equity premium puzzle in financial markets.[1] Further, experimental methods have been employed to examine whether psychological findings persist in market settings and how the structure of incentives in the environment interacts with human psychological pre-dispositions. Camerer (1987), for instance, investigated whether biases in probability judgments matter in market settings and found that they persist, although significantly reduced, in comparison to those observed outside of markets. Fehr and Gächter (2002), showed that when the members of a group of people invest individually in a mutually beneficial common good, many of them are willing to incur a non-negligible cost in order to punish free-riders. At the aggregate level, this "altruistic punishment" allows the group to maintain high levels of investment in the common good.

1.1.2 Computational social science

The use of computers and the internet has revolutionized the way in which we can study the relation between individual and aggregate behavior. Already in the 1970s, social scientists started to develop agent-based models to simulate complex social systems. Programming offered a precise language to represent individual behavior and the structure of the environment. In contrast to the colossal computational capacities of rational choice theory, the behavioral assumptions in this tradition of research are minimalistic. By and large, individual behavior is described by simple rules of thumb. However, there have been instances of modellers who have designed artificial agents that behave according to an actual psycho-

[1] The puzzle refers to the fact that stocks have in the long run a much higher return than bonds.

logical theory (Arthur, 1993; Spiliopoulos, 2013). Moveover, psychologists have collaborated with complex systems theorists in a quest to develop models of collective behavior with more realistic behavioral foundations (Galam and Moscovici, 1991; Lewenstein et al., 1992).

Recently, new computational tools, such as the possibility to run large-scale experiments (Salganik et al., 2006; Bond et al., 2012), and the availability of big datasets and theoretical developments such as network theory (see Lazer et al., 2009) have opened new research avenues at the intersection between psychology and the social sciences. For example, it is now possible to conduct controlled large-scale experiments over the internet and to collect precise behavioral data that can be used to reconstruct the link from the micro to the macro-level (Salganik et al., 2006). Further, using big datasets makes it possible to analyze how human behavior propagates from one individual to another in large social networks (Cacioppo et al., 2009; Fowler and Christakis, 2010).

It is now clear that social systems are characterized by peculiarities such as tipping points and latent behaviors that cannot be captured by rational choice models or naive statistical models. Nevertheless, the methods available to behavioral and social scientists today make it possible to uncover cognitive and social influence processes and to simulate collective behavior using computers.

1.2 Overview of models presented in this dissertation

In line with recent developments in behavioral economics and computational social science, this thesis pioneers an integrationist approach and seeks to reduce the conceptual distance between psychology and the social sciences. In the three papers included in this dissertation, I present novel psychologically inspired process models of individual behavior and examine their implications for aggregate behavior. These models leverage the potential of psychological methods and theories.

The work presented in chapter two departs from two milestone models of rational choice theory and integrates them using psychological linear estimation models. One the one hand, one-shot multi-attribute utility theory has been studied theoretically and empirically in management, marketing science and psychology. It the state-of-the-art model of choice in static environments. On the other hand, optimal stopping theory has been developed theoretically in statistics and economics. Further, human behavior in scenarios such as those described by the theory has been investigated in economics, psychology and marketing. We show that when human beliefs about the utility of alternatives can be represented as psychological linear models, the search problem has an intuitive and psychologically plausible solution that can be expressed in terms of sequential cost-benefit analysis. In our framework, the description of the environment is very similar to that of the probit and ordered probit-models which are widely used in econometrics. Thus, in future, the models we propose can be directly pitted against their naive statistical counterparts.

The model presented in chapter three brings together the experimental approach of social psychology with the modeling approach used within the field of complex systems. Previously, models of opinion formation have been based on simple behavioral assumptions in regard to belief revision without empirically testing them. In contrast, we conducted two controlled experiments and derived a process model directly from experimental observations. We combined cognitive and agent based modeling to describe the individual level processes and to track the behavior in the entire population. At the cognitive level, we designed a decision tree representing how people revise their opinion and confidence when exposed to the opinions of other individuals. Then, we scaled up the model to the aggregate level by allowing social interaction between the agents. This model clarifies the role of confidence in the formation of public opinion and can be used to investigate the impact of experts in shaping the public opinion.

Finally, the model presented in chapter four brings together the modeling approaches advanced in chapter two and three to investigate the markets for cultural products. We postulated a simple cognitive mechanism according to which decision makers search and choose alternatives. We assumed that decision makers with diverse yet correlated preferences search the alternatives in the order of popularity and stop after meeting an alternative with a utility higher than a satisficing threshold. The model can capture market-share inequality, final outcome unpredictability and the weak correlation between the quality and the popularity of the alternatives observed in markets for cultural products. In addition, it reproduces the predictable failure-unpredictable success pattern observed in real world cultural product markets and a recent large-scale experiment.

All three models provide examples of interdisciplinary work between psychology and the social sciences. I hope that the projects presented in this thesis will provide to historians of the future an account for the divide between psychology and the social sciences and more importantly will exemplify strategies to move beyond it.

1.3 References

Arthur, W. B. (1993). On designing economic agents that behave like human agents. *Journal of Evolutionary Economics*, 3(1):1–22.

Benartzi, S. and Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, 110(1):73–92.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., and Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295–298.

Cacioppo, J. T., Fowler, J. H., and Christakis, N. A. (2009). Alone in the crowd: the structure and spread of loneliness in a large social network. *Journal of Personality and Social Psychology*, 97(6):977.

Camerer, C. F. (1987). Do biases in probability judgment matter in markets? experimental evidence. *The American Economic Review*, pages 981–997.

Cyert, R. M. and March, J. G. (1963). *A behavioral theory of the firm.*

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, 51(4):380.

Erev, I. and Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88(4):848–881.

Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, 57(2):94.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868):137–140.

Fowler, J. H. and Christakis, N. A. (2010). Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences*, 107(12):5334–5338.

Galam, S. and Moscovici, S. (1991). Towards a theory of collective phenomena: consensus and attitude changes in groups. *European Journal of Social Psychology*, 21(1):49–74.

Gigerenzer, G. and Todd, P. M. (1999). *Simple heuristics that make us smart.* Oxford University Press.

Heukelom, F. (2009). *Kahnemann and Tversky and the making of behavioral economics.* Rozenberg Publishers.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291.

Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., et al. (2009). Life in the network: the coming age of computational social science. *Science (New York, NY)*, 323(5915):721.

Lewenstein, M., Nowak, A., and Latané, B. (1992). Statistical mechanics of social impact. *Physical Review A*, 45(2):763.

Papandreou, A. G. (1958). *Economics as a Science.* JB Lippincott.

Phillips, L. D. and Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of experimental psychology*, 72(3):346.

Salganik, M. J., Dodds, P. S., and Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762):854–856.

Schelling, T. C. (1978). *Micromotives and Macrobehavior.* WW Norton.

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, 25(3):289–310.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2):129.

Simon, H. A. (1998). Psychology and economics. In *The handbook of economic methodology*.

Spiliopoulos, L. (2013). Beyond fictitious play beliefs: Incorporating pattern recognition and similarity matching. *Games and Economic Behavior*, 81:69–85.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Chapter 2

Multi-attribute utility models as cognitive search engines

Pantelis P. Analytis, Amit Kothiyal and Konstantinos Katsikopoulos

Chapter Abstract

In optimal stopping problems, decision makers are assumed to search randomly to learn the utility of alternatives; in contrast, in one-shot multi-attribute utility optimization, decision makers are assumed to have perfect knowledge of utilities. We point out that these two contexts represent the boundaries of a continuum, of which the middle remains uncharted: How should people search intelligently when they possess imperfect information about the alternatives? We assume that decision makers first estimate the utility of each available alternative and then search the alternatives in order of their estimated utility until expected benefits are outweighed by search costs. We considered three well-known models for estimating utility: (i) a linear multi-attribute model, (ii) equal weighting of attributes, and (iii) a single-attribute heuristic. We used 12 real-world decision problems, ranging from consumer choice to industrial experimentation, to measure the performance of the three models. The full model (i) performed best on average but its simplifications (ii and iii) also had regions of superior performance. We explain the results by analyzing the impact of the models utility order and estimation error.[1]

[1]The authors would like to thank Alexia Delfino, Juliane Kämmer, Hrvoje Stojic, Peter Todd and Jack

2.1 Introduction

Do the following thought experiment: You are the human resources manager of a company and you are assigned the task of hiring a new employee. After advertising the position, you receive several dozen applications from candidates listing their skills and credentials (e.g. grade point average, work experience, programming skills). You can determine each candidate's potential only after inviting him or her for an interview. Let us assume that you can interview candidates sequentially and that you can decide to stop interviewing and hire a candidate after each interview. Crucially, making the effort to interview another candidate is costly. What is the best way to organize the interview process? First, you need to decide the order in which you will be inviting candidates. Then, after each interview you need to decide whether to make an offer to one of the interviewed candidates, thus stopping your search. The first problem is an ordering problem and the second a stopping problem.[2]

Clearly, if you could perfectly estimate the potential of the candidates on the basis of their credentials you could with certainty invite the best one for an interview. Then your dilemma would be reduced to a single choice, which has been extensively studied across the behavioral sciences (e.g. Keeney and Raiffa, 1993). On the other hand, if the credentials were not at all informative, you would have to invite people at random, and your problem would reduce to an optimal stopping problem. Such models have been developed formally in statistics (DeGroot, 1970) and economics (Stigler, 1961; Lippman and McCall, 1976) and human behavior in them has been tested empirically in psychology (Rapoport and Tversky, 1966; Lee, 2006), economics (Schotter and Braunstein, 1981; Hey, 1987; Sonnemans, 1998) and marketing (Zwick et al., 2003). Intuitively, most everyday decision-making problems lie between these two extreme cases; in reality, the attributes of the alternatives can be used to predict their utility but only imperfectly. There often remains some amount of uncertainty that cannot

Willis for their remarks on earlier versions of the paper and the members of the ABC group for their feedback during internal presentations. The authors would also like to thank Anita Todd for editing the manuscript.

[2]The essay has been published under the same title in Judgment and Decision Making in September 2014. PPA conceived the idea. PPA, AK and KK developed the theory. PPA analyzed the data. PPA, AK and KK wrote the paper.

be explained by the attributes.

There has been some work on ordered search: In a seminal paper, Weitzman (1979) put forward a general model for ordering alternatives and terminating search in search problems with recall, where decision makers initially have partial information about the alternatives but learn their true utility after paying a search cost. Although Weitzman's results readily generalize to scenarios with multi-attribute alternatives, only a recent study by Dzyabura (2013) considered explicitly how sequential search can be guided by a multi-attribute utility model. Earlier, Roberts and Lattin (1991) presented a model of consideration set formation in which the decision makers include alternatives in their consideration set guided by a compensatory multi-attribute utility model. In essence, Roberts and Lattin's model can be seen as an ordered search model, in which the number of alternatives that will be searched has to be decided once and for all before any search is performed.[3] However, the authors did not connect their results to search theory. Further, Moorthy et al. (1997) employed Weitzman's model to derive predictions about the length of search in consumer choice. They further specified the original model and assumed that decision makers are uncertain about a brand's utility. The consumer beliefs in their model are probabilistic, and more experienced consumers are better at differentiating brands in terms of utility. Moorthy et al. tested their predictions on data from the car market. Last, in economics, several papers started from the assumption that agents search the alternatives in an externally imposed (Arbatskaya, 2007) or subjectively defined (Bagwell and Ramey, 1994; Armstrong et al., 2009) order and studied the aggregate market behavior.

In this paper we show that when the decision makers' preferences can be described by a linear utility model, the problems introduced by Weitzman have an intuitive and psychologically plausible solution. Returning to our example, we will analytically show that the optimal policy is to follow the estimated utility order prescribed by your subjective utility

[3]This was the case in the first search models that were introduced in economics (e.g. Stigler, 1961). Those models were later called fixed-sample-size models. Morgan and Manning (1985) discussed in detail the points of divergence between fixed-sample-size models and sequential sampling models, in which the decision maker can stop search at any search step.

model; then stop when the expected return from seeing one more candidate for the job turns negative. In essence, the utility models play the role of cognitive search engines, generating the order in which alternatives are examined. We formally develop this approach and apply it to three models that have been studied extensively in the field of judgment and decision making: (i) multi-attribute linear utility, (ii) equal weighting of attributes and (iii) a single-attribute heuristic. The simpler models (ii) and (iii) have been shown to perform well under some conditions (e.g. scarce information available for calibrating models) in one-shot choice problems (Barron, 1987; Gigerenzer et al., 1999; Fasolo et al., 2007; Katsikopoulos, 2011). We then compare the performance of the models in 12 real-world environments ranging from consumer choice to industrial experimentation and examine how the models' expected utility order and estimation error influence their performance and length of search.

Conceptually, our approach illustrates that optimal stopping problems assuming random search and one-shot choice problems are the boundary cases of an ordered search problem with imperfect information. It provides a formal framework within which the assumptions made by ordered search models, such as those proposed by Bagwell and Ramey (1994), Moorthy et al. (1997) and Armstrong et al. (2009), can be further clarified. In practice, our approach extends discrete choice models by specifying the exact search process. It is a plausible alternative to Roberts and Lattin's (1991) theory of consideration set formation and it further advances our understanding of decision-making in environments with rank-ordered alternatives.

In what follows, as in the model presented by Weitzman (1979), we focus on a scenario in which decision makers learn the exact utility of an alternative after sampling it and can always choose alternatives that they have sampled in the past. In Section 2 we develop a formal framework of optimal ordering and stopping. In Section 3 we test three models in 12 real-world environments ranging from consumer choice to industrial experimentation and examine the ordering and estimation error components of the models. In Section 4 we discuss the possibility of connecting our findings to other work, assess the conceptual implications of our approach, and finally discuss the limitations and possible extensions of our framework.

2.2 The theoretical framework

2.2.1 The environment

There are n alternatives $A_1, ..., A_n$. Each alternative A_i, $i \in \{1, ..., n\}$ is associated with a vector of attributes $\mathbf{a_i} = (a_{i1}, ..., a_{ik})$ and a utility u_i of choosing it. The u_is are unknown but the $\mathbf{a_i}$s are known to the decision maker. The decision maker estimates u_i by $f(\mathbf{a_i})$. We assume that the estimation errors, ϵ_i, such that $u_i = f(\mathbf{a_i}) + \epsilon_i$, are iid Gaussian with mean μ and standard deviation σ. We call this equation the decision maker's subjective model. For each alternative, the decision maker can only learn the utility u_i by sampling the alternative and paying a cost c. To choose an alternative, the decision maker can sample as many items as desired. If the decision maker searches for k items, the cost to be paid is kc and the decision maker will choose out of these k the alternative with the highest utility.

2.2.2 The optimal strategy

Let S denote the set of alternatives already searched and $\bar{S}$ denote the set of alternatives not searched yet. That is,

$$S \cup \bar{S} = \{A_1, ..., A_n\}, \quad S \cap \bar{S} = \emptyset$$

The decision maker's problem is to determine the search order and the stopping rule. Let the variable y denote the maximum utility that the decision maker can obtain from the alternatives in S, $y = max_{AkS} u_k$, where A_k belongs to S. If the decision maker sampled just one more item A_k before stopping search then the subjective expected gain (i.e. the increase in utility minus cost) is (probabilities and expectations below are based on the decision maker's subjective model):

$$\begin{aligned} R(A_k) = P(u_k > y) \times E(u_k - y - c|u_k > y) - P(u_k \leq y) \times c \\ = P(u_k > y) \times E(u_k - y|u_k > y) - c. \quad (2.1) \end{aligned}$$

It is intuitive that the decision maker should keep on sampling as long as there exists an alternative $A_k \in \bar{S}$ such that its subjective expected gain $R(A_k) > 0$ [if all $R(A_k) < 0$, search should be stopped]. Given this, the decision maker should sample the alternative A_k that achieves the maximum subjective gain $R(A_k)$. It turns out that to maximize $R(A_k)$, it suffices to select the alternative with the highest $f(\mathbf{a_i})$ (for a proof, see Appendix 1):

Result 1. *If for two alternatives A_i, $A_j \in \bar{S}$, $f(\mathbf{a_i}) > f(\mathbf{a_j})$ then $R(A_i) > R(A_j)$.*

This result says that if the decision maker decides to sample one more item before terminating the search, then the choice should be the one that has maximum unconditional expectation $E(u_i) = f(\mathbf{a_i})$. This suggests the following policy:

Selection rule: Order the alternatives based on their unconditional expectation. Select the items for sampling in this order.

Stopping rule: If at any stage subjective expected gain is negative, terminate the search.

Note that the stopping rule can be applied only if the standard deviation of the estimation error σ is estimated. On the other hand, for the selection rule to be applied, only the parameters of the multi-attribute function $f(\mathbf{a_i})$ need to be estimated. For $\sigma = 0$ the decision-making problem reduces to a single choice. For $c = 0$ the decision maker searches all the alternatives. Note that for $c = 0$ and when there are only two alternatives the model reduces to the probit model.[4] If $\sigma = \infty$, $P(u_i > y)$ is equal to 0.5.

2.2.3 Subjective models

In our framework, the order of the alternatives A_i is based on their subjective utility $f(\mathbf{a_i})$. Thus, the decision maker's eventual success depends upon the multi-attribute utility function

[4] Then, for two alternatives A_i and A_j the probability that the alternative A_i will be chosen can be written as $\Phi(\frac{f(\mathbf{a_i})-f(\mathbf{a_j})}{\sqrt{(2)}\sigma})$

he or she uses to estimate utilities and generate an order of the alternatives. We present three psychologically plausible and widely used multi-attribute utility functions:

1. **Multi-attribute linear utility (MLU)**: $f((a_{i_1}, ..., a_{i_m})) = \sum_j \beta_j a_{i_j}$

MLU is one of the cornerstone models in research on multi-attribute decision making (Keeney and Raiffa, 1993). It is also the model used to derive consumer preferences in conjoint analysis surveys in marketing. MLU is the equivalent of multi-linear regression, which has been widely studied as a model of inductive inference in multi-cue learning (Hammond et al., 1964).

2. **Equal-weighted linear utility (EW)**: $f((a_{i_1}, ..., a_{i_m})) = \sum_j a_{i_j}$, where all the attributes $a_{i_1}, ..., a_{i_m}$ are normalized and brought to the same scale.

EW is a special case of MLU where all decision weights β_js are equal. It was originally proposed as an alternative to multi-linear regression by Dawes and Corrigan (1974).

3. **Single-attribute utility (SA)**: $f((a_{i_1}, ..., a_{i_m})) = a_{i_j}$, where a_{i_j} has the highest ecological validity among $\{a_{i_1}, ..., a_{i_m}\}$ as expressed by Kendall's tau non-parametric correlation.

Versions of the SA model have been studied by Hogarth and Karelaia (2005a). The SA model is akin to the lexicographic heuristic (Payne et al., 1993; Kohli and Jedidi, 2007) and the take-the-best heuristic (Gigerenzer and Goldstein, 1996; Katsikopoulos et al., 2010). However, SA resolves ties between alternatives by choosing at random, whereas the lexicographic heuristic and take-the-best examine additional attributes.

Note that the three presented models can also be seen as cardinal estimation models, where the criterion value corresponds to the utility. Davis-Stober et al. (2010b) have pointed out that the notation of utility and estimation models can be used interchangeably.

2.2.4 An example

We illustrate how the different components of equation 1 play out with a concrete example (see Figure 2.1). Consider a scenario where a decision maker is searching in an online store to buy a single album of an up-and-coming music band she just heard about on the radio. The band has produced three albums so far. The decision maker can learn the exact utility of an album by listening to its songs. Her subjective beliefs are described by the SA model. As

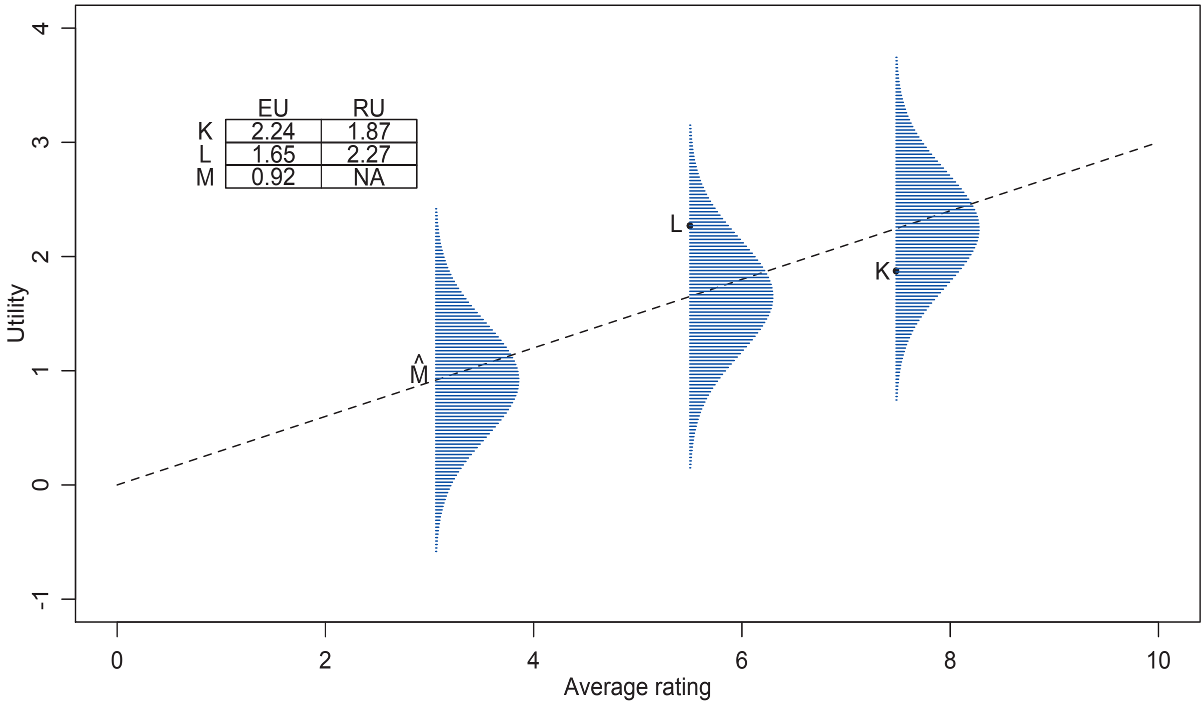


Figure 2.1: An example of the decision-making process. There are three alternatives available in the market represented by the points K, L and M on the plot. The average rating a_i (x axis) maps to the subjective utility (y axis) of the alternatives for the decision maker but with some error. As represented by the straight line, the decision maker believes that the expected utility of the alternatives can be estimated by $u_i = 0.3 * a_i$ where a_i is the average rating. As represented by the bell-shaped curves, the decision maker believes that the estimation error ϵ_i of her model is iid Gaussian with mean $\mu = 0$ and standard deviation $\sigma = 0.5$. The inset table shows the expected utility (EU) and realized utility (RU) of the alternatives. The decision maker will first sample the alternative with the highest attribute value and then at each step decide whether to sample the next alternative. For $c = 0.05$, for example, the decision maker will sample the alternatives K and L and choose the alternative L.

represented by the straight line in Figure 2.1, the decision maker believes that the expected utility of an alternative can be estimated by $u_i = 0.3 * a_i$, where a_i is the average rating of the album by other users of the site. As represented by the bell-shaped curves, she believes that the estimation error ϵ_i of her model is iid Gaussian with mean $\mu = 0$ and standard deviation $\sigma = 0.5$. The decision maker first samples album K, which has the highest expected utility. She finds out that the utility of album K is 1.87, which is slightly less than its expected utility. Then she has to decide whether it is worthwhile to examine the album L. Following equation 1 the expected returns from sampling album L can be written as $P(u_L > u_K) \times E(u_L - u_K | u_L > u_K) - c$. $P(u_L > u_K) = 0.33$, $E(u_L - u_K | u_L > u_K) = 0.329$

and their the product equals 0.109.[5] Thus, the decision maker will examine the second album if the cost of search is lower than 0.109; otherwise she will stop search, choose album K, and never learn the actual utility of album L. Let us assume that the cost is 0.05. Then, the decision maker samples album L. She finds out that the utility of L for her is 2.27, which is higher than her expectation and the utility of album K. Thus, L replaces K as the sampled album with the highest utility (y in equation 1). Now the returns from sampling album M can be written $P(u_M > u_L) \times E(u_M - u_L | u_M > u_L) - c$. $P(u_M > u_L) = 0.003$ and $E(u_M - u_L | u_M > u_L) = 0.152$. The product of these two parts equals 0.0005. Thus the overall return is negative and the decision maker will stop search after sampling album L and choose it. She will never learn the realized utility of album M.

2.3 Results

We applied the three models to study guided search in real-world problems. We examined the performance of the models in 12 data sets ranging from consumer choice to industrial experimentation. All the datasets included a variable with positive and more-is-better valence, which we treated as the utility. All the attribute values were normalized to a 0-1 scale to implement the EW model. When a correlation between an attribute and the criterion value in the test data set was positive, we converted the lowest attribute value to 0 and the highest to 1. When a correlation was found to be negative we converted the highest attribute value to 0 and the lowest to 1. We used ordinary least squares to calibrate the parameters β_j of the three models. For each of the models, we estimated the standard deviation of the error component σ using the standard error of the corresponding linear regression model $\hat{\sigma} = \sqrt{\frac{\sum_{t=0}^{T}(y-\hat{y})^2}{T}}$ where T is the number of alternatives in the training set and $\hat{y}$ the estimate of the linear regression model with parameters β_j.

[5]Note that $\sigma = 0.5$ appears in both of these two terms. If the decision maker believed $\sigma = 1$ the $P(u_L > u_K)$ would increase to 0.413 and $E(u_L - u_K | u_L > u_K)$ to 0.723 and the total benefits would be 0.299, almost three times higher than in the case where $\sigma = 0.5$.

2.3.1 Real world environments

Environment	No. A	No. a	R^2	$\lvert max(\rho_{ua_i})\rvert$	$\lvert\overline{\rho}_{ua_i}\rvert$	$\lvert\overline{\rho}_{aiaj}\rvert$
Beer aroma*	23	7	0.39	0.35	0.23	0.43
Cheese taste	30	3	0.65	0.76	0.67	0.65
CPU efficiency*	209	6	0.87	0.86	0.64	0.51
F. lamp lifetime*	14	3	0.57	0.33	0.24	0.41
Machine productivity	40	5	0.56	0.71	0.21	0.07
Octane quality	82	4	0.91	0.87	0.63	0.51
Olive oil quality	16	4	0.65	0.51	0.36	0.15
Potato taste	32	4	0.35	0.35	0.28	0.11
Red wine quality*	1599	11	0.36	0.48	0.19	0.2
Seed yield	30	3	0.70	0.81	0.36	0.18
Tea quality	64	5	0.39	0.43	0.21	0.21
White wine quality*	4898	11	0.28	0.44	0.15	0.17

Table 2.1: The environment names refer to the variable that was assumed to be the utility in that environment. Asterisks indicate that the source data set was composed of several alternatives that were available in the market at the point of data collection. All the remaining data sets represent the results of controlled experiments. No. A refers to the number of alternatives and No. a to the number of attributes. The notation $\lvert max(\rho_{ua_i})\rvert$ indicates the strongest Pearson correlation between the utility variable and the attributes, $\lvert\overline{\rho}_{ua_i}\rvert$ the mean Pearson correlation between the utility and the attributes, and $\lvert\overline{\rho}_{aiaj}\rvert$ the mean intercorrelation between the attributes. CPU stands for central processing unit and f. lamp for fluorescent lamp.

The 12 datasets analyzed are freely available from online databanks. We provide an additional reference list of the publications in which the datasets were originally reported in the supplementary material. As reported in Table 2.1, the datasets are characterized by a diverse number of attributes, alternatives, and intercorrelations between the attributes. We normalized the utility variable, setting the utility of the alternative with the lowest value equal to 0 and that with the highest value equal to 1. This transformation was necessary in order to achieve comparability across the datasets. The remaining variables are the attributes that were used to predict the alternatives' utility. In a few cases we excluded variables that were not suitable as attributes in a choice problem. In the white wine quality and the red wine quality environments we reduced the number of alternatives to 200 by choosing once at random from the initial dataset. This transformation was implemented to generate a plausible decision-making ecology. The ecological characteristics of the resulting environments were very similar to those of the complete environments reported in Table 2.1.[6] We used the same

[6]for red wine quality : $R^2 = 0.3$, $\lvert max(\rho_{ua_i})\rvert = 0.43$, $\lvert\overline{\rho}_{ua_i}\rvert = 0.16$, $\lvert\overline{\rho}_{aiaj}\rvert = 0.20$; for white wine quality

	Average performance		
Cost	MLU	EW	SA
$1/2^3$	0.606	0.594	0.561
$1/2^4$	0.715	0.696	0.681
$1/2^5$	0.783	0.775	0.761
$1/2^6$	0.824	0.822	0.816
$1/2^7$	0.857	0.851	0.856
$1/2^8$	0.885	0.878	0.881
$1/2^9$	0.902	0.899	0.901
$1/2^{10}$	0.912	0.912	0.911

Table 2.2: The average performance of the models across the 12 environments in 8 cost conditions. Multi-attribute linear utility (MLU) performed best but as the cost decreased performance differences with equal-weighted linear utility (EW) and single-attribute utility (SA) attenuated.

subset of 200 alternatives consistently in all the analyses reported in this paper.

2.3.2 Performance

The performance of a model depends (1) on the order it generates, which determines how quickly it discovers high quality alternatives, and (2) on the estimated standard deviation of the error component $\hat{\sigma}$, which influences the subjective probability that the utility of the next alternative to be sampled will be higher than the utility of the best alternative discovered up to that point. Clearly, when $\hat{\sigma}$ increases, the subjective probability and consequently the expected returns from further search also increase. Finally, the model performance depends (3) on the cost of search in the environment. Factor 3 is an environmental factor, while factors 1 and 2 reflect how well a model captures the properties of the environment.

At first, we manipulated the cost of search and pitted the models against each other in eight different cost conditions. For the purposes of our simulation we adapted the technique of cross-validation to a search problem. We fixed the parameters β_j and σ corresponding to each model in half of the data set (training set) and evaluated the performance of the models in the remaining half (test set). This process was repeated 10.000 times in total. For each repetition the alternatives that were part of the training and test sets were drawn at random from the entire data set. As a result, the maximum utility, which could be achieved by all

:$R^2 = 0.35$, $|max(\rho_{ua_i})| = 0.48$, $|\overline{\rho}_{ua_i}| = 0.16$, $|\overline{\rho}_{aiaj}| = 0.19$

models in each of the repetitions, differed slightly depending on the utilities of the alternatives that were part of the training and test sets. The models first sampled the alternative that they estimated to have the highest expected utility. Then each model decided according to equation 1 whether to proceed to search the alternative with the second highest expected utility, then the third, and so forth. When a model stopped search, the alternative with the highest utility among the alternatives searched up to that point was chosen. The measure of performance was the average utility achieved by a model in 10.000 repetitions. In real life, this corresponds to the average performance of 10.000 decision makers, with randomly sampled experiences, who all followed the optimal policy. This measure can also be seen as an approximation of the expected payoff for a single decision maker.

Average results from the 12 environments for all eight cost conditions are presented in Table 2.2 and the results from each of the 12 environments individually are presented in Figure 2.2. The best performing model was MLU. The performance differences were largest for high search costs and they gradually attenuated as the costs decreased. Clearly, a high cost implies that the models are likely to search only a few alternatives at the beginning of the order, and the models may have different orders. As the costs decrease, all the models search further down the expected utility line and are more likely to discover the high-quality alternatives in the environment. As a result, the performance differences diminish.

There is much variability in performance in individual environments, where in several cases the simpler models outperformed MLU for the entire cost range. For example, EW performed best for most of the cost conditions in the beer aroma, cheese taste, CPU efficiency, olive oil quality, and potato taste environments, while SA performed best in the red wine quality, tea quality, and octane quality environments. Thus, although MLU performed better on average, it clearly outperformed the simpler models only in the remaining four environments. The better performance of MLU at the aggregate level was mainly driven by its superior performance in the white wine quality environment. As stressed earlier, the model-specific factors that may influence the models' performance are on the one hand the order generated by each model and on the other hand the estimated standard deviation. In the rest of the Results section we decouple the roles of these two factors.

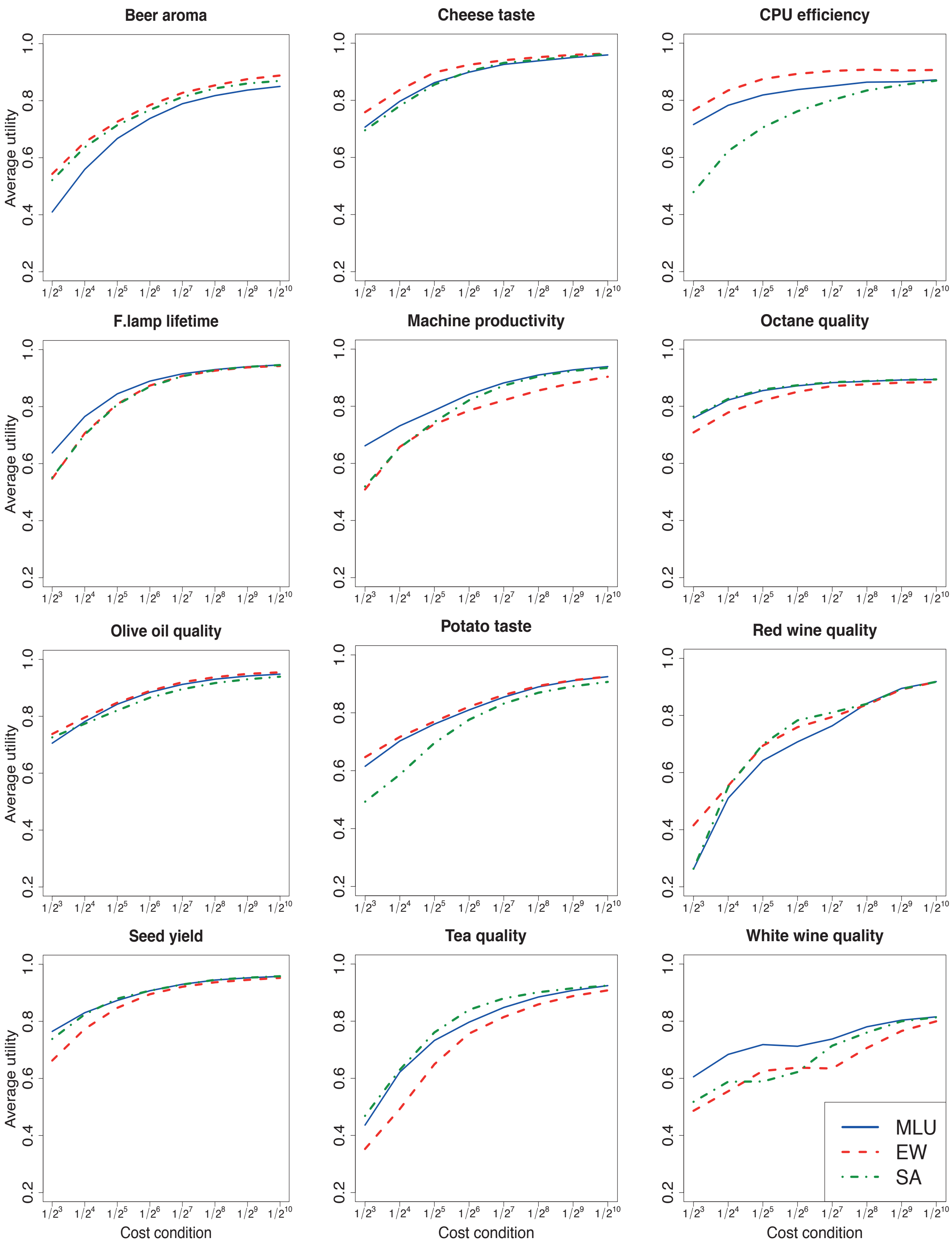


Figure 2.2: The average utility achieved by the subjective utility models as a function of search cost. Overall, multi-attribute linear utility (MLU) performed best, but equal-weighted linear utility (EW) and single-attribute utility (SA) also had regions of superior performance. The difference between the models gets smaller as the cost of search decreases. CPU stands for central processing unit and f. lamp for fluorescent lamp.

2.3.3 Impact of search order on performance and search length

To disentangle the impact of the search order from that of the estimated standard deviation of the error component $\hat{\sigma}$, we first examined the average return to search (as measured by the average utility of the best searched alternative), achieved by each multi-attribute utility model, for all possible search lengths k, without implementing any search cost. Moreover, we compared these results to random search, which corresponds to the assumption made in most optimal-stopping studies. As we did for the full task, we cross-validated the models on half of the data points and we repeated our simulations 10.000 times.

As shown in Figure 2.3, the largest difference between the subjective models and random search is found in environments with high R^2 of the best fitting linear regression. In the environments with the highest R^2, such as CPU efficiency and octane quality, the best solution was almost always located in one of the first search trials by all the multi-attribute models. In contrast, in environments with low R^2 such as beer aroma, potato taste, and fluorescent lamp lifetime, the margin is smaller and in some cases random search performed almost as well as the multi-attribute models.

In all environments there is a close correspondence between the average return to search as a function of the search length and the performance of the model in the full task, as depicted in Figure 2.2. For most costs, the model with the highest average return to search is also the best performing model in the full task. In four of the five environments in which EW performed best in the full task (beer aroma, cheese taste, CPU efficiency, potato taste) it also performed best for most of the search lengths. Similarly, in the three environments where SA performed best in the full task (red wine quality, tea quality and octane quality) it also performed as well as or better than the other two models for most search lengths in the search task. Overall, there are a few cases, such as the tea quality environment, where different models lead to best performance for different search lengths k.

Note that the factors that influence the performance of a model directly influence the length of search, as observed in Figure 2.4. According to the theory presented in section 2.2, the utility of the best alternative secured so far, y, and the probability $P(u_k > y)$, that the utility u_k of the next alternative down the expected utility line A_k will be higher than y should be inversely related. This suggests that models that place high-utility alternatives earlier in the search order will be characterized by lower returns to search and should, ceteris paribus, search less. Indeed, this is what we observe in Figures 2.3 and 2.4. However, to fully understand the performance and search length of the models we also need to come to grips with the role played by the standard deviation of the error component σ. The next section delves into that exactly.

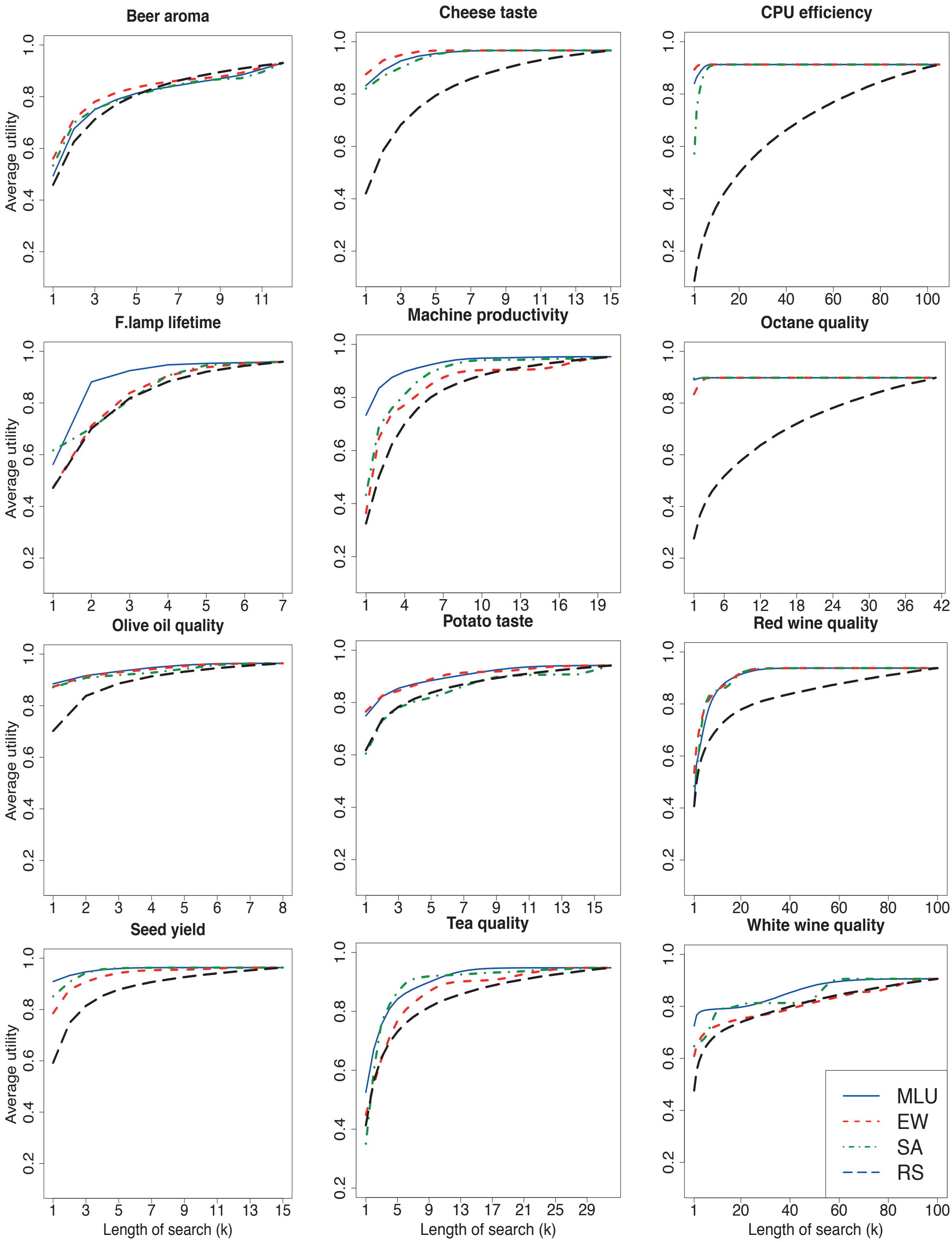


Figure 2.3: The average utility of the best explored alternative after a search of length k. In environments with a high R^2, such as central processing unit (CPU) efficiency and octane quality, the best alternatives are located early in the search and the differences between these strategies and random search are the largest. In most cases there is a close correspondence between model performance in the full task and model performance in mere search; contrast with Figure 2.2. MLU stands for multi-attribute linear utility, EW for equal-weighted linear utility, and SA for single-attribute utility.

2.3.4 The role of the standard deviation σ

The second factor that has an impact on the performance of the model and the length of search is the standard deviation of the error component σ. To assure good performance the models' error component should correspond to the unexplained uncertainty in the environment. For estimated $\hat{\sigma} = 0$ the model would deem that the first alternative is also the best one. For $\hat{\sigma} = \infty$ the model would calculate $P(u_k > y) = 0.5$ (following equation 1). However, both these models would be unadaptive in environments in which the uncertainty that cannot be explained by $f(\mathbf{a_i})$ is moderate. A lower standard deviation of the error component $\hat{\sigma}$ implies a decrease in the subjective probability $P(u_k > y)$ and a decrease in the expected returns from finding a better alternative $E(u_k - y|u_k > y)$. Thus, ceteris paribus, it should lead to a shorter search.

	Average search length		
Cost	MLU	EW	SA
$1/2^3$	1.74	1.58	1.74
$1/2^4$	2.23	2.28	2.45
$1/2^5$	2.81	2.83	3.21
$1/2^6$	3.70	3.62	4.47
$1/2^7$	4.88	5.39	5.58
$1/2^8$	6.11	7.24	6.72
$1/2^9$	7.34	8.85	8.08
$1/2^{10}$	8.66	10.45	9.54

Table 2.3: The average length of search (number of alternatives sampled k) across environments for the three models for different costs of search. The equal-weighted linear utility (EW) and single-attribute utility (SA) models search more alternatives, on average, than multi-attribute linear utility (MLU). The differences in the length of search are more pronounced for low costs.

Within models, the estimated standard deviation of the error component $\hat{\sigma}$ in our simulations is inversely related to the accuracy of the estimates of the model in the training set in which the model was fitted. Consequently, in most cases a good ordering should be accompanied by a relatively low error component and in tandem they lead to a reduced length of search. Note, however, that there could be a discrepancy between the accuracy of the estimates of a model and the average returns to search. Remember that only the best alternative discovered so far counts for the decision maker in the search task.

Between models, the MLU tends to capture a larger proportion of the uncertainty in $\mathbf{a_i}$ ($\bar{\hat{\sigma}}_{mlu} = 0.166$) and usually has a slightly lower estimated standard deviation $\hat{\sigma}$ than EW and SA ($\bar{\hat{\sigma}}_{ew} = 0.191$ and $\bar{\hat{\sigma}}_{sa} = 0.191$). As reported in Figure 2.4, only in the olive oil quality and potato taste environments was a heuristic model (EW) found to have an estimated standard deviation of the error component lower than the full model (MLU).

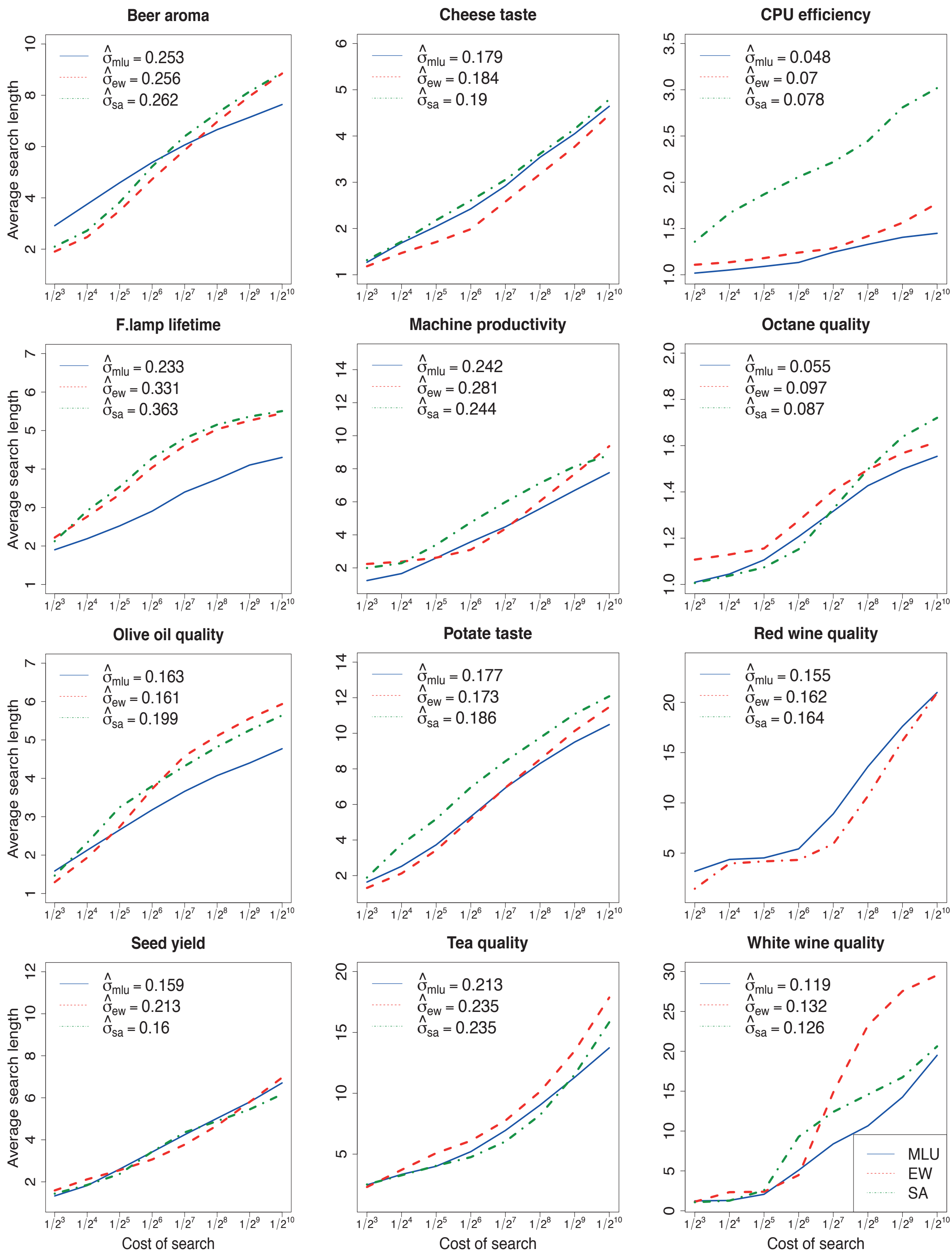


Figure 2.4: On the upper left side we present the average standard deviation of the error component of each of the models in that environment. The length of search is moderated by the best alternative discovered so far and the estimated deviation of the error component of the models. On average, multi-attribute linear utility (MLU) searches less than the equal-weighted linear utility (EW) and single-attribute utility (SA) models but there is some variability across environments.

As seen in Figures 2.2-4, the models that achieve higher average returns to search (Figure 2.3), search less (Figure 2.4, Table 2.3 for average results) and score better on the full task (Figure 2.2). The differences in the estimated standard deviation of the models $\hat{\sigma}$ are in most environments marginal. In half of the environments they correspond to performance differences in the average returns to search. This suggests that the error components of the models are sufficiently well calibrated to the unexplained uncertainty in the environment. There are a few cases where a different estimated standard deviation $\hat{\sigma}$ might lead the models to markedly different performance. For example, in the CPU efficiency environment the EW and MLU have similar returns to search. MLU has a much lower estimated standard deviation $\hat{\sigma}$, searches less than EW, and performs worse in the full task. This indicates that MLU may occasionally miss a very good alternative and it could benefit from searching more.

2.3.5 Paired-comparison results

So far, we have discussed how different properties of the search models influence their performance and we have looked at specific environments to understand how these factors play out in practice. We have illustrated that the most crucial factor is the search order in which the models sample the alternatives, although this has to be accompanied by a well-calibrated estimated standard deviation of the error component $\hat{\sigma}$. But is there a way to predict which strategy is most likely to be successful in a given environment? The conditions under which different linear or heuristic models perform well in choice and inference contexts have been thoroughly investigated (for a review see Katsikopoulos, 2011). Few studies generalize beyond binary choice to choice between several alternatives or among the entire data set; however, most of the existing literature has focused on binary choices. To answer our question we attempt to shed light on the connections between the novel task we presented and the well-studied paired-comparison task.

Thus, we examined whether there is a correspondence in performance between the binary choice and the search task. We compared the performance of the three models in the binary choice task. As before, we fixed the parameters corresponding to each model in half of the dataset and evaluated their performance in all possible binary choice tasks in the remaining half. The process was repeated in total 10.000 times.

	Subjective models		
Environment	MLU	EW	SA
Beer aroma	0.48	0.49	0.47
Cheese taste	0.79	0.79	0.76
CPU efficiency	0.83	0.85	0.80
F. lamp lifetime	0.67	0.54	0.55
Machine productivity	0.64	0.55	0.63
Octane quality	0.86	0.72	0.79
Olive oil quality	0.72	0.71	0.58
Potato taste	0.63	0.64	0.57
Red wine quality	0.63	0.60	0.63
Seed yield	0.78	0.64	0.77
Tea quality	0.69	0.62	0.63
White wine quality	0.65	0.60	0.66
Mean performance	0.70	0.65	0.65

Table 2.4: Accuracy of the three models on binary choices. Fifty percent of the dataset was used as a training set and all the possible pairs of the test set were used to evaluate the models. Multi-attribute linear utility (MLU) performed best, on average, followed by single-attribute utility (SA) and equal-weighted linear utility (EW) . MLU performed best in seven environments EW in four and SA in one.

As seen in Table 2.4, on average, MLU performed best, followed by SA and then EW. In the individual environments, the performance of the models varied significantly. There were four environments in which EW performed best and one in which SA did. In general, the performance in the binary choice task is a good proxy of performance in the full search task. The model that performed best in the paired-comparison task also performed best in 7 of the 12 environments in the search task for cost equal to $1/2^3$. The discrepancies observed can be attributed to the fact that in the full search task, the alternatives that are searched early on contribute disproportionally to the success of a model. In contrast, in binary choices all possible single choices in the data set contribute equally to the performance of the model. Hence, it is possible that the model that performs best in the search task does not maintain that superior performance in the binary choice task, and vice versa.

2.4 Discussion

2.4.1 Conceptual implications of our framework

Deterministic optimization and search models: A possible compromise

In optimization problems, widely studied in economics, decision makers can determine the alternative that maximizes their utility. This vision of decision making contrasts with search models in which the decision makers sample alternatives at random and stop search after encountering a good enough alterative (e.g. Simon, 1955; Chow et al., 1971; Caplin et al., 2011). Random sampling may lead to violations of the revealed preference principle and unpredictability in regard to the choices of individual decision makers. Ordered search models provide a possible compromise between these two approaches. Decision makers have a well-defined utility function before the search starts. However, as long as there is some uncertainty about the exact utility of the alternatives, it may pay to sample some of them to learn their utility. In our model, the initial preferences guide the search process but are also subject to revision when the true utility of the sampled alternatives is revealed. For an external observer, such as a firm or a market analyst, ordered search is more predictable, at the level of an individual decision maker, than random search. In ordered search, if the model of the decision maker and the actual utility of the alternatives are known, the external observer could also predict the final decisions made, as well as the preference reversals that would occur along the way.

Evaluating the hypotheses of ordered search models

Employing our framework, we can examine the conditions under which the assumptions postulated in previous ordered search models hold true. Bagwell and Ramey (1994) suggested that decision makers may use a simple rule of thumb and first consider buying products from firms that advertise more.[7] This corresponds to an SA model where the amount of advertising is the most informative cue. This, indeed, is plausible in some cases. However, the SA model implies that when more informative attributes are available, the amount of advertisement may be completely ignored as an attribute. Armstrong et al. (2009) assumed that consumers search products according to an abstract attribute called prominence and suggested that firms might be willing to invest to achieve prominence and improve their search order. Our framework allows us to estimate the exact impact of an intervention in the attributes of a product in terms of the firm's prominence in the market. Finally, Moorthy et al. (1997) suggested that the expectation of the utility of a brand's products is normally distributed. In addition, they suggested that more experienced consumers are better able to differentiate between products. These assumptions are both in line with our framework.

2.4.2 Applied implications of our framework

Sequential search and consideration set formation

Our modeling approach suggests that only the alternatives that have been sampled by the decision makers stand a chance of being selected. Similarly, several marketing scientists have advanced choice models in which the decision makers first restrict their attention to a subset of the alternative set — commonly called the consideration set. The decision makers then examine the alternatives of this subset more closely and finally choose one of the alternatives in it (e.g. Wright and Barbour, 1977; Shocker et al., 1991; Gilbride and Allenby, 2004; Moe, 2006). Such models have often been found to outperform, in fitting and prediction, discrete choice models in which the decision makers are assumed to consider all the alternatives (e.g. Gilbride and Allenby, 2004). Our approach shares some of its assumptions with a popular model of consideration set formation put forward by Roberts and Lattin (1991). In their model, decision makers, whose pay-off function is described by a compensatory multi-attribute utility model with an additional error term, decide which alternatives to include in

[7] A similar argument for the role of advertisement in the domain of choice has been made by Goldstein and Gigerenzer (2002)

their consideration set. Similar to our approach, the decision makers examine the alternatives in the order of decreasing expected utilities predicted by their utility model, paying a fixed cost for every new alternative they place in their consideration set. In contrast to our model, the decision makers do not learn the exact utility of the alternatives immediately after paying the cost. Instead, they learn all the utilities of the alternatives in the consideration set right before they choose among them. Our model can be seen as the sequential search counterpart to Roberts and Lattin theory of consideration set formation. Inversely, Roberts and Lattin's model can also be understood as an ordered search model in which the search length has to be decided at the outset. The exact domain of application of sequential search and fixed-sample-size models of consideration-set formation may depend on the exact characteristics of the decision making context and it should be the subject of further empirical investigation in the future.

Parallels to online ranking schemes

There are clear-cut parallels between the search approach we present here and the methods used by commercial search engines and recommendation systems to pre-rank the alternatives for Internet users. Such ranking schemes implement a valence function that maps the attributes of the alternatives to the relevance or utility for the user (Burges et al., 2005; Hüllermeier et al., 2008; Yaman et al., 2011). The valence functions underlying the ranking schemes are trained with clickstream decision data or other information that can reveal the preferences of the user (Sarwar et al., 2000). Then, similar to what our theory suggests, the ranking schemes present the alternatives in decreasing order of relevance or utility. This approach is implemented without invoking a formal decision-making theory that predicts how decision makers will choose on the basis of the presented rank order. If the goal of the ranking-scheme engineers is to increase the utility derived by the users, a formal decision-making theory, such as search theory, might further inform the development of ranking-schemes as well as the techniques used to train their valence functions. Indeed, there have been recent papers in machine learning about why the design of ranking techniques can benefit from taking into account how people actually decide in rank-ordered environments (Agichtein et al., 2006; Chapelle et al., 2009). We take a step in that direction and illustrate the role of cost of search and the uncertainty in the environment in the search process.

2.4.3 Connections to prescriptive and descriptive decision making

From choice to search

For the paired-comparison problem, where the task is to choose one of two alternatives, the accuracy of heuristics such as EW and SA has been analyzed and compared to that of the full linear model in numerous studies. Until now, a thread of the existing literature in judgment and decision making has examined binary choices in environments with binary or continuous attributes without making any assumptions about the mapping from an alternative's attributes to its utility (e.g. Katsikopoulos et al., 2010) or when the mapping is characterized by noise (Hogarth and Karelaia, 2005a, 2006; Rieskamp and Otto, 2006; Davis-Stober et al., 2010a,b). A second thread of the literature has focused on environments with binary or continuous attributes, where a mapping between the attributes and the utility exists but decision makers have imprecise knowledge about the attribute weights (Johnson and Payne, 1985; Martignon and Hoffrage, 2002; Hogarth and Karelaia, 2005b; Baucells et al., 2008; Katsikopoulos, 2013).

Given the observed correspondence in choice task and search task results, the findings of the first thread of studies in binary choice tasks may also generalize to the search task. Overall, these studies have found that there are no large performance differences between the heuristics and the full model and that heuristics can outperform the full model under the appropriate conditions. Both EW and SA fare especially well in out-of-sample prediction (Einhorn and Hogarth, 1975; Hogarth and Karelaia, 2005b; Katsikopoulos et al., 2010). The SA model tends to perform well when a simply or cumulatively dominating alternative is present (Baucells et al., 2008; Şimşek, 2013; Katsikopoulos et al., 2014), or when there exist high correlations between the single attribute and all other attributes (Hogarth and Karelaia, 2005a; Davis-Stober et al., 2010a). EW tends to perform well when the variability in cue validities is small or when there are high intercorrelations between all the attributes (Einhorn and Hogarth, 1975; Wainer, 1976). In the environments that we studied we also found support for some of these findings. For example in the four environments, in which EW performed best in binary choice the difference between $|max(\rho_{ua_i})|$ and $|\overline{\rho}_{ua_i}|$ was small.

Although our first results suggest a relation between the search task and the choice task, additional research is required in the future to establish this relation and to identify the conditions under which differences in the relative performance of models in the two tasks are to be expected. A further follow-up to our study would be to examine when decision makers choose a certain strategy (Bröder, 2003) and how the decision makers learn to act adaptively and to select a strategy that performs well in a given environment (e.g. Rieskamp and Otto, 2006).

Psychological plausibility of the stopping rule

So far we have shown that the proposed stopping policy is optimal. But is it psychologically plausible? When sampling from a known distribution with or without recall an optimally acting decision maker should always stop right after encountering an alternative with a value higher than the optimal threshold. Several variations of such optimal threshold problems have been studied extensively in psychology and economics. We have identified five studies reporting results on experiments in which subjects sampled from a known distribution with recall (Rapoport and Tversky, 1970; Schotter and Braunstein, 1981; Hey, 1987; Kogut, 1990; Sonnemans, 1998). In sum, a moderate proportion of the participants in these studies behaved in a manner consistent with the optimal stopping rule. Common discrepancies from the optimal strategy included stopping too early and exercising recall. Nonetheless, the researchers found that the average performance of the participants was near optimal. Hey (1982) and Sonnemans (1998) reported that many subjects used heuristic strategies that appeared consistent with the optimal rule and led to near optimal performance. In the stopping policy we have presented, the decision maker should at every search step reevaluate the returns from sampling the next alternative in line. This task may appear demanding in relation to the optimal threshold rule. However, it has the same structure as a simplified version of the signal-detection problem (Green and Swets, 1966) — a task in which humans are known to perform fairly well. Thus, we believe that the stopping policy could be psychologically plausible as such or it could be well approximated by clever heuristic algorithms. Clearly, the human behavior in the task has to be investigated experimentally in the near future.

Alternative search algorithms

We showed that under the homoscedasticity assumption embedded in many linear models and the subjective linear utility models outlined in this paper, the intuitive policy of searching the alternatives in the order of their subjective expected utility is optimal. In addition, in this special case our policy coincides with another simple algorithm called *directed cognition*, which has been proposed by Gabaix et al. (2006). This algorithm searches myopically, as if the next step of the search process was also the last one. When the assumption of the homogeneously distributed error component ϵ does not hold, the behavior suggested by our ordering policy diverges from Weitzman's (1979) optimal solution and from Gabaix et al.'s (2006) directed cognition algorithm. Gabaix et al. (2006) showed that even in simple cases of indexable problems like those discussed by Weitzman (1979) and Gittins et al. (1989), the decision makers are unlikely to follow the optimal algorithm and instead decide in line with the directed cognition algorithm. In the future, as in Gabaix et al. (2006), one could examine experimental scenarios in which the behavior prescribed by the three discussed algorithms diverges to evaluate their psychological plausibility in different decision-making environments.

2.4.4 Extensions and limitations

The current model could be extended to scenarios with different search costs for different alternatives or when the decision maker can choose more than one alternative. In addition, it is possible to account for contexts where new alternatives become available at a later point in time. The decision maker can construct a search order for the new alternatives and consider if it is worthwhile to sample some of them, examining first the alternative with the highest expected utility.

In our model we have assumed that decision makers have direct and free access to the attributes all at once. We did not discuss cases where the decision makers have to pay a search cost to learn additional attributes before moving forward to examine further alternatives. A model of this kind with random sampling is presented by Lim et al. (2006) and has been investigated empirically by Bearden and Connolly (2007). Further, as in search problems, we have assumed that the decision maker learns the exact utility of an alternative after paying a search cost and examining it. However, there are many dynamic decision-making contexts where the cost is internally defined as an opportunity cost when consuming an inferior alternative (Nelson, 1970). In these environments information acquisition can be inherently noisy and decision makers may want to sample the alternatives repeatedly. Such decision-making contexts are commonly referred to as multi-armed bandits. In fact, when we change the costly sampling assumption to repeated sampling, in which the final payoff is the sum of the experienced utilities, our framework turns into a multi-armed bandit with contextual information. This type of bandit framework has been receiving increasing attention in recent years in machine learning (Pavlidis et al., 2008; Li et al., 2010).

So far we have postulated that decision makers have stable preferences and an accurate error estimate of their model throughout the entire search process. This strong assumption may hold precisely or it may approximate the truth in many decision-making environments, especially when decision makers have long experience testing alternatives. However, in some environments decision makers may not know the utility weights but rather learn them along the way as they examine new alternatives. This approach has been followed by Dzyabura (2013), who assumed that decision makers update the estimates of the weights and the search order after each new alternative they examine. Clearly, when decision makers learn their preferences the ordering and stopping rules derived in our paper are not guaranteed to be optimal. A fully rational policy would have to anticipate the future evolution of decision makers' preferences and then build any beliefs about preference change into the search and stopping rules. Even in simplified scenarios this approach is known to be computationally intractable. In similar preference-learning problems encountered in machine learning, greedy heuristics are implemented instead to balance preference learning and exploitation (Brochu et al., 2007, 2010). Another approach would be to compare, for any given search length, the performance of alternative, optimal in some respect, active learning algorithms (Fedorov, 1972; Sugiyama and Nakajima, 2009).

2.4.5 Conclusion

In two recent publications Luan et al. (2011; 2014) stressed the need to integrate decision-making theories in psychology and illustrate how apparently disparate models share common conceptual ground. In the same vein, we argued that choice and search problems, which until now have been studied separately, are the boundary cases of a broader decision-making problem. We showed how three choice models that have been extensively studied in the field of judgment and decision making can guide the search for good alternatives and we formulated their corresponding optimal stopping rule. Then, we compared the performance of the models in 12 real-world environments ranging from consumer choice to industrial experimentation and illustrated how each models' expected utility ordering and estimation error influence its' performance and length of search. As in previous model comparisons, in one-shot choice problems we found that heuristic linear models performed on average close to a multi-attribute linear utility model. Moreover, in individual environments the heuristic models often outperformed the full model. To further understand when such results are to be expected we examined the relationship between the search problem and the well-studied binary choice problem. We found that in most cases the models that performed well on the binary choice task also did so in the search task. This suggests that previous findings on the ecological rationality of choice and inference strategies are also relevant to the search task. Finally, we discussed the connections of our model to the existing literature and suggested possible paths for future research.

2.5 References

Agichtein, E., Brill, E., and Dumais, S. (2006). Improving web search ranking by incorporating user behavior information. In *Proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval*, pages 19–26. ACM.

Arbatskaya, M. (2007). Ordered search. *The RAND Journal of Economics*, 38(1):119–126.

Armstrong, M., Vickers, J., and Zhou, J. (2009). Prominence and consumer search. *The RAND Journal of Economics*, 40(2):209–233.

Bagwell, K. and Ramey, G. (1994). Coordination economies, advertising, and search behavior in retail markets. *The American Economic Review*, 84(3):498–517.

Barron, F. H. (1987). Influence of missing attributes on selecting a best multiattribute alternative. *Decision Sciences*, 18(2):194–205.

Baucells, M., Carrasco, J. A., and Hogarth, R. M. (2008). Cumulative dominance and heuristic performance in binary multiattribute choice. *Operations Research*, 56(5):1289–1304.

Bearden, J. N. and Connolly, T. (2007). Multi-attribute sequential search. *Organizational Behavior and Human Decision Processes*, 103(1):147–158.

Brochu, E., Cora, V. M., and De Freitas, N. (2010). A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*.

Brochu, E., De Freitas, N., and Ghosh, A. (2007). Active preference learning with discrete choice data. In *Advances in neural information processing systems*, pages 409–416.

Bröder, A. (2003). Decision making with the" adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(4):611–625.

Burges, C., Shaked, T., Renshaw, E., Lazier, A., Deeds, M., Hamilton, N., and Hullender, G. (2005). Learning to rank using gradient descent. In *Proceedings of the 22nd international conference on Machine learning*, pages 89–96. ACM.

Caplin, A., Dean, M., and Martin, D. (2011). Search and satisficing. *The American Economic Review*, 101(7):2899–2922.

Chapelle, O., Metlzer, D., Zhang, Y., and Grinspan, P. (2009). Expected reciprocal rank for graded relevance. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 621–630. ACM.

Chow, Y. S., Robbins, H., and Siegmund, D. (1971). *Great expectations: The theory of optimal stopping*. Houghton Mifflin, Boston.

Davis-Stober, C. P., Dana, J., and Budescu, D. V. (2010a). A constrained linear estimator for multiple regression. *Psychometrika*, 75(3):521–541.

Davis-Stober, C. P., Dana, J., and Budescu, D. V. (2010b). Why recognition is rational: Optimality results on single-variable decision rules. *Judgment and Decision Making*, 5(4):216–229.

Dawes, R. M. and Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, 81(2):95.

DeGroot, M. (1970). *Optimal statistical decisions*. New York [etc.]: McGraw-Hill, New York.

Dzyabura, D. (2013). The role of changing utility in product search. *Available at SSRN: http://ssrn.com/abstract=2202904 or http://dx.doi.org/10.2139/ssrn.2202904*.

Einhorn, H. J. and Hogarth, R. M. (1975). Unit weighting schemes for decision making. *Organizational Behavior and Human Performance*, 13(2):171–192.

Fasolo, B., McClelland, G. H., and Todd, P. M. (2007). Escaping the tyranny of choice: When fewer attributes make choice easier. *Marketing Theory*, 7(1):13–26.

Fedorov, V. V. (1972). *Theory of optimal experiments*. Academic Press Inc., New York.

Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *The American Economic Review*, 96(4):1043–1068.

Gigerenzer, G. and Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological review*, 103(4):650–669.

Gigerenzer, G., Todd, P. M., and the ABC Research Group (1999). *Simple heuristics that make us smart*. Oxford University Press, USA, New York.

Gilbride, T. J. and Allenby, G. M. (2004). A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science*, 23(3):391–406.

Gittins, J., Glazebrook, K., and Weber, R. (1989). *Front matter*. Wiley Online Library.

Goldstein, D. G. and Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109(1):75–90.

Green, D. M. and Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley New York, New York.

Hammond, K. R., Hursch, C. J., and Todd, F. J. (1964). Analyzing the components of clinical inference. *Psychological Review*, 71(6):438–456.

Hey, J. D. (1982). Search for rules for search. *Journal of Economic Behavior & Organization*, 3(1):65–81.

Hey, J. D. (1987). Still searching. *Journal of Economic Behavior & Organization*, 8(1):137–144.

Hogarth, R. M. and Karelaia, N. (2005a). Ignoring information in binary choice with continuous variables: When is less more? *Journal of Mathematical Psychology*, 49(2):115–124.

Hogarth, R. M. and Karelaia, N. (2005b). Simple models for multiattribute choice with many alternatives: When it does and does not pay to face trade-offs with binary attributes. *Management Science*, 51(12):1860–1872.

Hogarth, R. M. and Karelaia, N. (2006). Regions of rationality: Maps for bounded agents. *Decision Analysis*, 3(3):124–144.

Hüllermeier, E., Fürnkranz, J., Cheng, W., and Brinker, K. (2008). Label ranking by learning pairwise preferences. *Artificial Intelligence*, 172(16):1897–1916.

Johnson, E. J. and Payne, J. W. (1985). Effort and accuracy in choice. *Management Science*, 31(4):395–414.

Katsikopoulos, K. V. (2011). Psychological heuristics for making inferences: Definition, performance, and the emerging theory and practice. *Decision Analysis*, 8(1):10–29.

Katsikopoulos, K. V. (2013). Why do simple heuristics perform well in choices with binary attributes? *Decision Analysis*, 10(4):327–340.

Katsikopoulos, K. V., Egozcue, M., and Garcia, L. F. (2014). Cumulative dominance in multi-attribute choice: Benefits and limits. *EURO Journal on Decision Processes*, 2:1–11.

Katsikopoulos, K. V., Schooler, L. J., and Hertwig, R. (2010). The robust beauty of ordinary information. *Psychological Review*, 117(4):1259–1266.

Keeney, R. L. and Raiffa, H. (1993). *Decisions with multiple objectives: Preferences and value trade-offs*. Cambridge University Press.

Kogut, C. A. (1990). Consumer search behavior and sunk costs. *Journal of Economic Behavior & Organization*, 14(3):381–392.

Kohli, R. and Jedidi, K. (2007). Representation and inference of lexicographic preference models and their variants. *Marketing Science*, 26(3):380–399.

Lee, M. D. (2006). A hierarchical bayesian model of human decision-making on an optimal stopping problem. *Cognitive Science*, 30(3):1–26.

Li, L., Chu, W., Langford, J., and Schapire, R. E. (2010). A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World Wide Web*, pages 661–670. ACM.

Lim, C., Bearden, J. N., and Smith, J. C. (2006). Sequential search with multiattribute options. *Decision Analysis*, 3(1):3–15.

Lippman, S. A. and McCall, J. (1976). The economics of job search: A survey. *Economic Inquiry*, 14(2):155–189.

Luan, S., Schooler, L., and Gigerenzer, G. (2014). From perception to preference and on to inference: An approach-avoidance analysis of thresholds. *Psychological Review*, 121(3):501.

Luan, S., Schooler, L. J., and Gigerenzer, G. (2011). A signal-detection analysis of fast-and-frugal trees. *Psychological Review*, 118(2):316.

Martignon, L. and Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, 52(1):29–71.

Moe, W. W. (2006). An empirical two-stage choice model with varying decision rules applied to internet clickstream data. *Journal of Marketing Research*, 43(4):680–692.

Moorthy, S., Ratchford, B. T., and Talukdar, D. (1997). Consumer information search revisited: Theory and empirical analysis. *Journal of Consumer Research*, 23(4):263–277.

Morgan, P. and Manning, R. (1985). Optimal search. *Econometrica: Journal of the Econometric Society*, 53(4):923–944.

Nelson, P. (1970). Information and consumer behavior. *The Journal of Political Economy*, 78(2):311–329.

Pavlidis, N. G., Tasoulis, D. K., and Hand, D. J. (2008). Simulation studies of multi-armed bandits with covariates. In *Proceedings of the tenth international conference on computer modeling and simulation*, pages 493–498. IEEE Computer Society.

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press, Cambridge.

Rapoport, A. and Tversky, A. (1966). Cost and accessibility of offers as determinants of optional stopping. *Psychonomic Science*, 4(4):145–146.

Rapoport, A. and Tversky, A. (1970). Choice behavior in an optional stopping task. *Organizational Behavior and Human Performance*, 5(2):105–120.

Rieskamp, J. and Otto, P. E. (2006). Ssl: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135:207–236.

Roberts, J. H. and Lattin, J. M. (1991). Development and testing of a model of consideration set composition. *Journal of Marketing Research*, 28(4):429–440.

Sarwar, B., Karypis, G., Konstan, J., and Riedl, J. (2000). Analysis of recommendation algorithms for e-commerce. In *Proceedings of the 2nd ACM conference on electronic commerce*, pages 158–167. ACM.

Schotter, A. and Braunstein, Y. M. (1981). Economic search: An experimental study. *Economic Inquiry*, 19(1):1–25.

Shocker, A. D., Ben-Akiva, M., Boccara, B., and Nedungadi, P. (1991). Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions. *Marketing letters*, 2(3):181–197.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.

Şimşek, Ö. (2013). Linear decision rule as aspiration for simple decision heuristics. In *Advances in Neural Information Processing Systems*, pages 2904–2912.

Sonnemans, J. (1998). Strategies of search. *Journal of Economic Behavior & Organization*, 35(3):309–332.

Stigler, G. J. (1961). The economics of information. *The Journal of Political Economy*, 69(3):213–225.

Sugiyama, M. and Nakajima, S. (2009). Pool-based active learning in approximate linear regression. *Machine Learning*, 75(3):249–274.

Wainer, H. (1976). Estimating coefficients in linear models: It don't make no nevermind. *Psychological Bulletin*, 83(2):213.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, 47(3):641–654.

Wright, P. and Barbour, F. (1977). Phased decision strategies: Sequels to an initial screening. In *Multiple criteria decision making : North Holland TIMS studies in the management sciences*, pages 91–109. North-Holland, Amsterdam.

Yaman, F., Walsh, T. J., Littman, M. L., and Desjardins, M. (2011). Democratic approximation of lexicographic preference models. *Artificial Intelligence*, 175(7):1290–1307.

Zwick, R., Rapoport, A., Lo, A. K. C., and Muthukrishnan, A. (2003). Consumer sequential search: Not enough or too much? *Marketing Science*, 22(4):503–519.

2.6 Appendix 1: Proof of result 1

First notice that for any alternative A_k

$$
\begin{aligned}
R(A_k) = P(u_k > y) \times E(u_k - y | u_k > y) - c = \int_y^\infty (u-y) dF_K(u) - c \\
= \int_0^\infty u dF_K(u) - c
\end{aligned}
$$

For alternative A_i and A_j let $\alpha = f(\mathbf{a_i}) - f(\mathbf{a_j})$. Assuming $\alpha > 0$, we will show that $R(A_i) > R(A_j)$. The distribution of A_i is the same as the distribution of $A_j + \alpha$; that is, $F_i(u) = F_j(u - \alpha)\ \forall u \in \Re$.

$$
\begin{aligned}
R(A_i) + c = \int_0^\infty u dF_i(u) = \int_0^\infty u dF_j(u-\alpha) \\
= \int_{-\alpha}^\infty (\alpha + v) dF_j(v) \text{ (by change of variable } v = u + \alpha) \\
= \int_{-\alpha}^0 (\alpha + v) dF_j(v) + \int_0^\infty (\alpha + v) dF_j(v) > \int_0^\infty (\alpha + v) dF_j(v) \\
> \int_0^\infty v dF_j(v) = R(A_j) + c
\end{aligned}
$$

Q.E.D

Chapter 3

Social influence and the collective dynamics of opinion formation

Mehdi Moussaïd, Juliane E. Kämmer, Pantelis P. Analytis and Hansjörg Neth

Chapter Abstract

Social influence is the process by which individuals adapt their opinion, revise their beliefs, or change their behavior as a result of social interactions with other people. In our strongly interconnected society, social influence plays a prominent role in many self-organized phenomena such as herding in cultural markets, the spread of ideas and innovations, and the amplification of fears during epidemics. Yet, the mechanisms of opinion formation remain poorly understood, and existing physics-based models lack systematic empirical validation. Here, we report two controlled experiments showing how participants answering factual questions revise their initial judgments after being exposed to the opinion and confidence level of others. Based on the observation of 59 experimental subjects exposed to peer-opinion for 15 different items, we draw an influence map that describes the strength of peer influence during interactions. A simple process model derived from our observations demonstrates how opinions in a group of interacting people can converge or split over repeated interactions. In particular, we identify two major attractors of opinion: (i) *the expert effect*, induced by the presence of a highly confident individual in the group, and (ii) *the majority effect*, caused by

the presence of a critical mass of laypeople sharing similar opinions. Additional simulations reveal the existence of a tipping point at which one attractor will dominate over the other, driving collective opinion in a given direction. These findings have implications for understanding the mechanisms of public opinion formation and managing conflicting situations in which self-confident and better informed minorities challenge the views of a large uninformed majority.[1][2]

[1]The authors are grateful to Gudrun Rauwolf and Tor Nielsen for their inspiring feedbacks and participation in the work. The authors thank Jack Soll for sharing with us the experimental data from ref. [30], and two anonymous reviewers for their constructive comments. The authors thank Anita Todd for language editing.

[2]This essay has been published under the same title in PLOS ONE in November 2013. MM JEK PPA HN conceived and analyzed the experiments. MM JEK PPA HN performed the experiments. MM JEK PPA analyzed the data. MM JEK PPA HN wrote the paper.

3.1 Introduction

In many social and biological systems, individuals rely on the observation of others to adapt their behaviors, revise their judgments, or make decisions (Bikhchandani et al., 1992; Camazine et al., 2001; Moussaid et al., 2009; Couzin et al., 2011). In human populations, the access to social information has been greatly facilitated by the ongoing growth of communication technology. In fact, people are constantly exposed to a steady flow of opinions, advice and judgments of others about political ideas, new technologies, or commercial products (Wu and Huberman, 2007). When facing the opinions of peers on a given issue, people tend to filter and integrate the social information they receive and adjust their own beliefs accordingly (Gigerenzer et al., 1999; Yaniv, 2004). At the scale of a group, repeated local influences among group members may give rise to complex patterns of opinion dynamics such as consensus formation, polarization, or fragmentation (Schelling, 1978; Isenberg, 1986; Galam and Moscovici, 1991; Mäs et al., 2010). For example, it has been shown that people sharing similar extreme opinions, such as racial prejudices, tend to strengthen their judgment and confidence after interacting with one another (Myers and Bishop, 1970). Similar mechanisms of opinion dynamics can take place in a variety of social contexts, such as within a group of friends exchanging opinions about their willingness to get vaccinated against influenza (Funk et al., 2009, 2010). At even larger scales, local influences among friends, family members, or coworkers often combined with the global effects of mass media constitute a major mechanism driving opinion formation during elections, shaping cultural markets (Salganik et al., 2006), producing amplification or attenuation of risk perceptions (Slovic, 1987; Kasperson et al., 1988), and shaping public opinion about social issues, such as atomic energy or climate change (Latane, 1981).

Given the remarkably large scope of social phenomena that are shaped by social influence and opinion dynamics, it is surprising that the behavioral mechanisms underlying these processes remain poorly understood. Important issues remain open: How do people adjust their judgment during social interactions? What are the underlying heuristics of opinion adaptation?

And how do these local influences eventually generate global patterns of opinion change? Much of the existing modeling work about opinion dynamics has been addressed from a physics-based point of view, where the basic mechanisms of social influence are derived from analogies with physical systems, in particular with spin systems (Galam, 1997; Schweitzer and Hołyst, 2000; Sznajd-Weron and Sznajd, 2000; Lorenz, 2007; Castellano et al., 2009). The wide variety of existing models assumes that individuals hold binary or continuous opinion values (usually lying between -1 and 1), which are updated over repeated interactions among neighboring agents. Different models assume different rules of opinion adaptation, such as imitation (Liggett, 1987), averaging over people with similar opinions (Deffuant et al., 2000; Hegselmann and Krause, 2002), following the majority (Galam, 2002), or more sophisticated equations (Sznajd-Weron and Sznajd, 2000; Mäs et al., 2010). Although informative as to the complex dynamics that can possibly emerge in a collective context, these simulation-based contributions share a common drawback: the absence of empirical verification of the models assumptions (Sobkowicz, 2009). Indeed, it is difficult to track and measure how opinions change under experimental conditions, as these changes depend on many social and psychological factors such as the personality of the individuals, their confidence level, their credibility, their social status, or their persuasive power (Latane, 1981). In other disciplines such as psychology and cognitive science, laboratory experiments have been conducted to study how people integrate feedback from other individuals to revise their initial answers to factual questions (Yaniv, 2004; Soll and Larrick, 2009; Lorenz et al., 2011). However, the findings on local rules of opinion adaptation have not yet been used to study the *collective* dynamics of the system, and it remains unclear how social influence plays out in larger scale social contexts over time (Mason et al., 2007).

The present work draws upon experimental methods inspired by social psychology and theoretical concepts of complex systems typical of statistical physics. First, we conducted controlled experiments to describe the micro-level mechanisms of social influence, that is, how individuals revise their initial beliefs after being exposed to the opinion of another person. Then, we elaborated an individual-based model of social influence, which served to investigate the collective dynamics of the system. In a first experiment (see Section 3.2), 52 participants were instructed to answer a series of 32 general knowledge questions and evaluate their confidence level on a scale ranging from 1 (*very unsure*) to 6 (*very sure*). This baseline experiment was used to characterize the initial configuration of the system before any social influence occurs. In a second experimental session, 59 participants answered 15 questions in the same way but were then exposed to the estimate and confidence level of another participant (henceforth referred to as "feedback") and asked to revise their initial answer. This procedure renders opinion changes traceable, and the effects of social influence *measureable* at the individual level. Moreover, changes in confidence were tracked as well, by asking participants to evaluate their confidence level before and after social influence. Despite empirical evidence suggesting that changes of opinion and confidence are intimately related (Lorenz et al., 2011), and theoretical work emphasizing the important role of inflexible, highly confident agents (Galam and Jacobs, 2007; Martins and Galam, 2013), this aspect of social influence remains poorly understood. Following the methods of existing experiments, we deliberately asked neutral, general knowledge questions, which allows capturing the mechanisms of opinion adaptation while controlling its emotional impact (Yaniv, 2004; Soll and Larrick, 2009). By exploring a simple model derived from our observations, we demonstrate that the collective dynamics of opinion formation in large groups of people are driven by two major "attractors of opinion": (i) the presence of a highly confident individual and (ii) the presence of clusters of low-confidence individuals sharing a similar opinion. In particular, we show that a critical amount of approximately 15 % of experts is necessary to counteract the attractive effect of a large majority of lay individuals. As people are embedded in strongly connected social networks and permanently influence one another, these results constitute a first step toward a better understanding of the mechanisms of propagation, reinforcement, or polarization of ideas and attitudes in modern societies.

3.2 Experiments

3.2.1 Experimental design

The experimental part of the study consisted of two distinct experiments: one without social influence (Experiment 1) and one with (Experiment 2). In both experiments, participants entered the laboratory individually and were instructed to answer a series of factual questions displayed on a computer screen. All participants were naïve to the purpose of our experiments and received a flat fee of 8 Euro. In Experiment 1, a total of 52 participants ($M_{age} = 27$ years, $SD = 9$, 50 % females) responded to 32 general knowledge questions, which covered the areas sports, nature, geography and society/economy (8 per area; for a complete list of items see Table 3.1). The correct answers to the questions ranged from 100 to 999, which, however, was not known to the participants. Participants were instructed to respond as accurately as possible and to indicate their confidence on a 6-point Likert scale (1 *very unsure* to 6 *very sure*) after having given their spontaneous estimate. Questions were displayed one after the other on the computer screen, and a new question was given only after participants answered the current one. Participants were only informed about the correct answers to the questions after the end of the experiment and therefore could not figure out that the true values always lied in the interval [100 999]. The order of the questions was randomized for each participant. A correlation test of the accuracy of answers and the order of the questions yielded non-significant p-values for 90 % of participants with a probability $p > 0.05$, confirming the absence of any learning process over experimental rounds. After the end of the experiments, participants were paid, thanked and released.

In Experiment 1, participants were not exposed to the social influence of others. The 1664 data points (corresponding to 52 participants $\times$ 32 questions) were used to characterize the features of the initial environment, such as the distribution of answers and the analyses of the confidence levels shown in Figure 3.1, and as a pool of social influence for the second experiment. The same dataset was used to define the initial condition of the simulations presented in Figure 3.5. In Experiment 2, 59 participants ($M_{age} = 33$ years, $SD = 11$, 56 % females) responded to 15 of the 32 general knowledge questions used in Experiment 1 and indicated their confidence level. Experiment 2 was conducted under the same conditions as in Experiment 1 except that participants were informed that they would receive a feedback from another participant. After each question, the estimate and confidence level of another randomly selected participant from Experiment 1 were displayed on the computer screen, and participants were then asked for a revised estimate and corresponding confidence level. This second dataset made of $59 \times 15 = 885$ binary interactions was used to study the effects of social influence, from which we derived the results shown in Figures 3.2, 3.3 and 3.4. The full list of questions is available in Table 3.1.

Ethics Statement: The present study has been approved by the Ethics Committee of the Max Planck Institute for Human Development. All participants gave written and informed consent to the experimental procedure.

3.2.2 Experimental results

We first use the data from the first experiment to characterize the initial configuration of the system before any social influence occurs, that is, how opinions are initially distributed and how the accuracy and confidence of the answers are correlated with each other. As shown in the example in Figure 3.1A, the initial distribution of opinions has a lognormal shape, with a typical long tail indicating the significant presence of outliers. For each of 32 items we performed a Kolmogorov-Smirnov normality test of $\log(O_i)$, where O_i is the initial opinion of individual i. The test yielded p-values above .05 for 84 % of the items, indicating that the null hypothesis cannot be rejected at the 5 % significance level for these items. The remaining 16 % still had reasonably high p-values (always $> 10^{-3}$), suggesting that the initial opinions O_i indeed follow a lognormal distribution.

We also analyzed the correlation between the confidence level of the participants and the

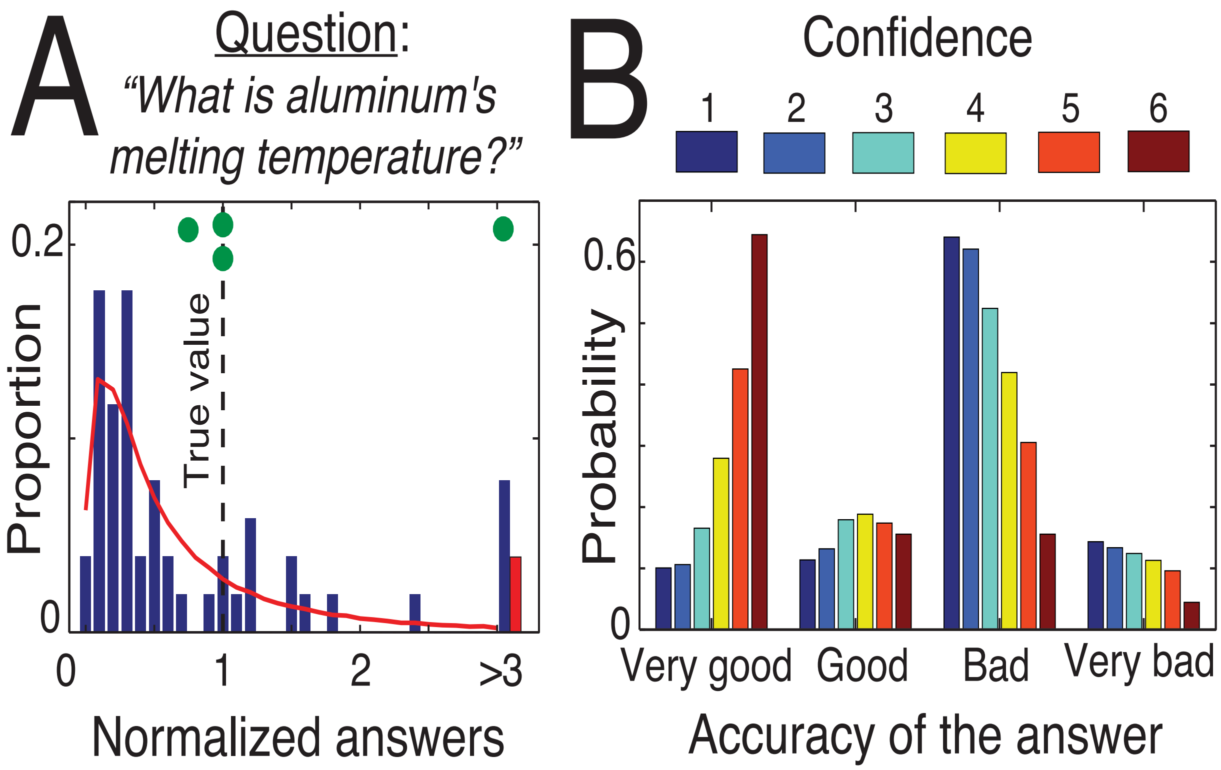


Figure 3.1: The initial configuration of the system in the absence of social influence. (A) Initial distribution of opinions for one representative example question (see Figure 3.7 for an overview of all 32 items). The normalized answer corresponds to the estimate of the participants divided by the true value (i.e., 660°C for this question). The red curve shows the best fit of a lognormal distribution. The green dots at the top indicate the location of estimates associated with high confidence levels ($C_i \geq 5$). One of them constitutes an outlier. (B) Accuracy of participants answers as a function of their confidence level, as determined from the complete dataset (32 items).

accuracy of their answer (Figure 3.1B). Interestingly, the confidence level is not such a reliable cue for accuracy (Nickerson, 1998). First, we found no significant correlation between an individual i's confidence level C_i and the quality of his or her answer (a correlation test between C_i and the error $Err(O_i) = |1 - O_i/T|$ where T is the true value yielded a coefficient of .03). Nevertheless, a trend can be highlighted by grouping the data into classes of error ranges: very good answers ($Err(O_i) \leq 0.1$), good answers ($0.1 < Err(O_i) \leq 0.3$), bad answers ($0.3 < Err(O_i) \leq 1$) and very bad answers ($Err(O_i) \geq 1$). As it can be seen from Figure 3.1B, only the maximum confidence level $C_i = 6$ is a relevant indicator of the quality of the answer, leading to a good or very good estimate in 80 % of the time. By contrast, lower confidence levels are less informative about accuracy. For instance, the second highest confidence value of $C_i = 5$ has a 39 % chance to correspond to a bad or very bad estimate. Similarly, a value of $C_i = 4$ is more likely to accompany a bad or very bad estimate (53 %) than a good or very good one (47 %). The lowest confidence values $C_i = 1$ and $C_i = 2$ do not differ from each other. Taking the revised estimates of Experiment 2 into account, we observe that the reliability of high confidence judgments is undermined by social influence (Lorenz et al., 2011). As shown in Figure 3.2B, the distribution of errors for very confident individuals ($C_i = 5$ or 6) becomes more noisy, widespread and clustered around certain values thus becoming less informative about accuracy after social influence.

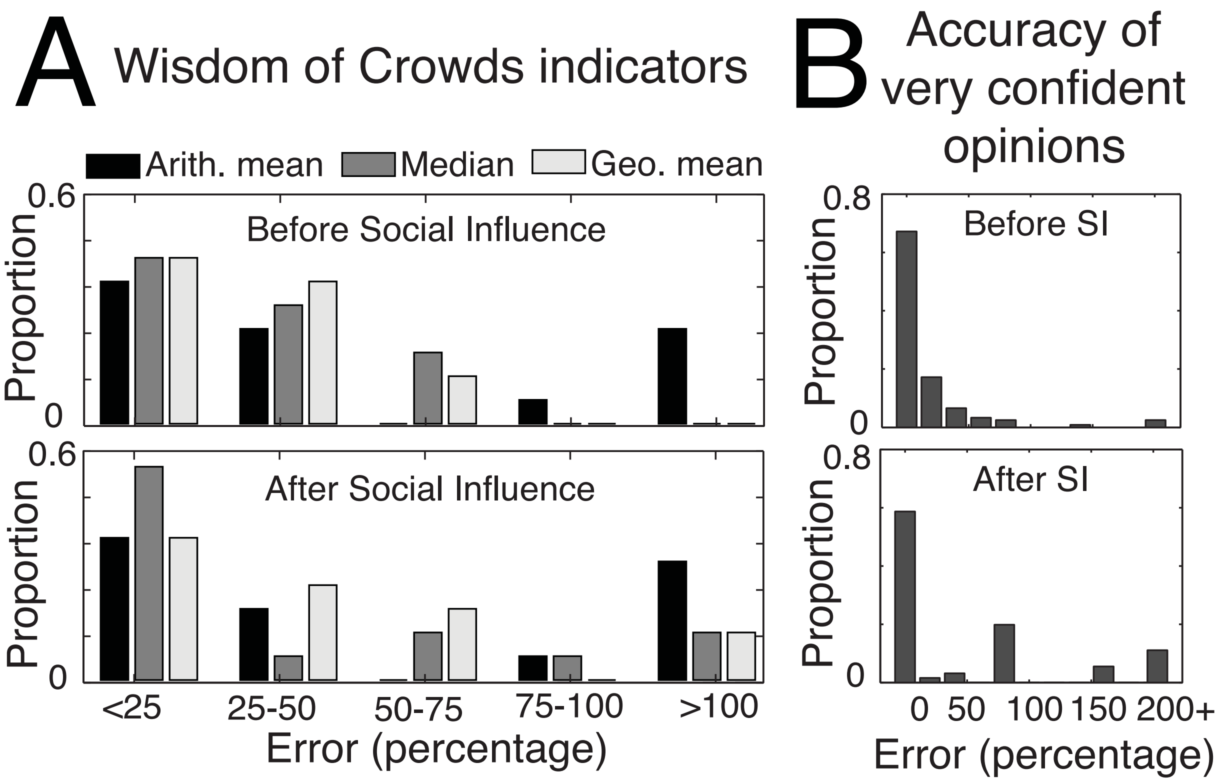


Figure 3.2: Effects of social influence on the wisdom of crowds (A), and the relevance of the confidence cue (B). The error is the deviation from the true value as a percentage. (A) Before any social influence occurs, the arithmetic (Arith.) mean is sensitive to single extreme opinions and does not appear as a relevant aggregating method. The median and geometric (Geo.) mean are more robust to outliers. When social influence occurs, however, the distributions are skewed to the right and the three indicators are more likely to generate high error values. (B) In the absence of social influence (SI), a clear and continuous trend is visible, where individuals with high confidence ($C_i \geq 5$) constitute a good indicator of the quality of the answer. When social influence is injected in the system, however, the distribution becomes noisier and less predictable. Overall, social influence generates unpredictability in the observed trends.

To explore the wisdom of crowds, we compared the accuracy of various aggregating methods before and after social influence occurred (Figure 3.2A). Our results agree with previous findings (Salganik et al., 2006; Lorenz et al., 2011). We find that the error distributions tend to become widespread, now covering a greater proportion of also high error values after social influence, regardless of the aggregating method.

Next, we focus on how people adjust their opinion after being informed about the opinion of another individual, which is the aim of Experiment 2. In agreement with previous studies (Yaniv, 2004; Soll and Larrick, 2009), our results show that two variables have an important influence on how the individual i revises his or her opinion when exposed to the opinion and confidence of another participant j: the difference in confidence values $\Delta C_{ij} = C_i - C_j$ and the normalized distance between opinions: $\Delta O_{ij} = |O_j - O_i|/O_i$, where O_j and C_j represent the opinion and confidence level of participant j, respectively (Yaniv, 2004).

To provide a visual, quantitative overview of the effects of social influence, we draw an influence map that illustrates the interplay of these two variables in the process of opinion adaptation (Figure 3.3). For the sake of simplicity, we distinguish three possible heuristics (Soll and Larrick, 2009):

1. *Keep initial opinion*, when individuals do not change their judgment after receiving a feedback, that is: $R_i = O_i$, where R_i is the revised opinion of participant i.

2. *Make a compromise*, when the revised opinion falls in between the initial opinion O_i and the feedback $O_j : min(O_i, O_j) < R_i < max(O_i, O_j)$.

3. *Adopt other opinion*, when an individual i adopts the partners opinion: $R_i = O_j$.

The influence map shows the heuristic that is used by the majority of people as ΔC_{ij} and ΔO_{ij} change (Figure 3.3A). Most of the data points (86 % of 885) are found for $-3 \leq \Delta C_{ij} \leq 3$ and $\Delta O_{ij} \leq 1.2$, which cover a large part of the influence map and seem to be reasonable ranges

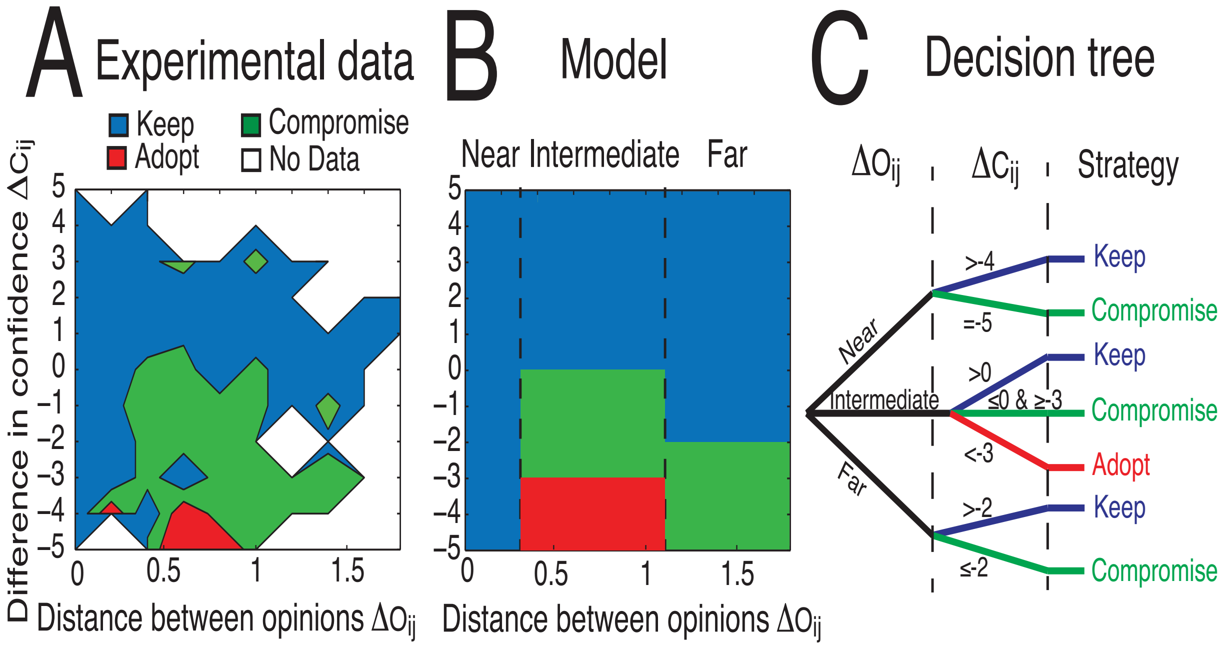


Figure 3.3: (A) The influence map extracted from our experimental data and (B) a simplified representation of it as implemented in the model. The color coding indicates the heuristic that is used by a majority of people, as a function of the difference in confidence $\Delta C_{ij} = C_i - C_j$ and the distance between the normalized opinions $\Delta O_{ij} = |O_j - O_i|/O_i$. Positive values of ΔC_{ij} indicate that the focus subject is more confident than the influencing individual (called feedback), whereas negative values indicate that the focus subject is less confident. White zones in (A) indicate the absence of sufficient data. Although the majority of people prefer to keep their initial opinion when they are more confident than their partner (i.e. the blue strategy dominates for $\Delta C_{ij} > 0$), a zone of strong influence is found at an intermediate distance with $\Delta C_{ij} < 0$. (C) The decision tree describing the decision process with three different outcome strategies. The individual first looks at the distance between opinions ΔO_{ij}, then looks at the difference of confidence ΔC_{ij}, and finally chooses a strategy accordingly.

being also encountered in real life situations. At the edge of the map, however, the results are more uncertain due to the scarcity of available data points. Figure 3.3A shows that the first and more conservative strategy tends to dominate the two others. In particular, the majority of people systematically keep their opinion when the value of ΔC_{ij} is positive, that is, when their own confidence exceeds their partners (Soll and Larrick, 2009). However, when their confidence level is equal or lower than their partners, individuals tend to adapt their opinion accordingly. Importantly, one can distinguish three zones in the influence map, according to the distance between estimates ΔO_{ij} (Figure 3.3B). First, when both individuals have a similar opinion ($\Delta O_{ij} < 0.3$), individuals tend to keep their initial judgment, irrespective of their partners confidence. Moreover, they also have a strong tendency to increase their confidence level (see Figure 3.4A indicating the changes in confidence). Therefore, we interpret this area of agreement as being a *confirmation zone*, where feedback tends to simultaneously reinforce initial opinions and increase an individuals confidence.

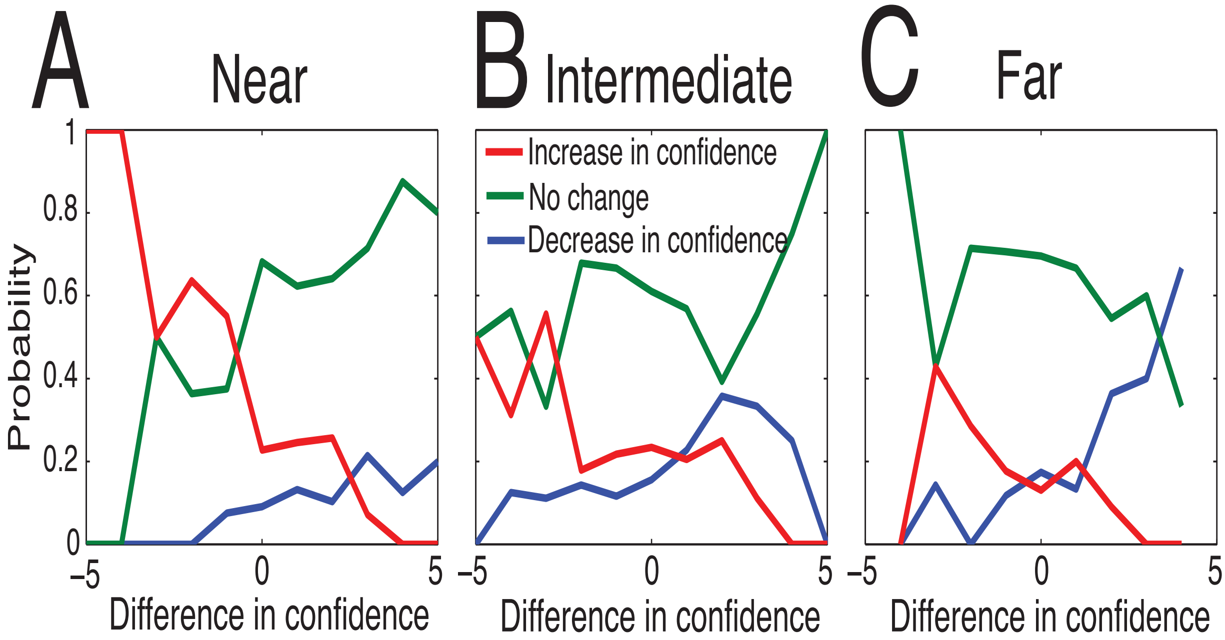


Figure 3.4: The probability of increasing (red), decreasing (blue), or maintaining (green) the confidence level after social influence. Changes in confidence are indicated according to the opinion distance classes as defined in the influence map (Figure 3.3): (A) near when $\Delta O_{ij} \leq 0.3$, (B) intermediate when $0.3 < \Delta O_{ij} \leq 1.1$, and (C) far when $\Delta O_{ij} > 1.1$. A tendency to increase confidence is visible in the near and intermediate zones when participants interact with a more confident subject. Confidence can also decrease in the far zone, when $\Delta C_{ij} \geq 4$.

It turns out that feedback has the strongest influence at intermediate levels of disagreement, when $0.3 < \Delta O_{ij} < 1.1$. In this zone, the *"compromise"* heuristic is selected by most people when $-3 \leq \Delta C_{ij} \leq 0$, and the *"adoption"* heuristic appears for lower values of ΔC_{ij}. We call this the *influence zone*, where social influence is strongest. Here, the others opinion differs sufficiently from the initial opinion to trigger a revision but is still not far enough away to be completely ignored. In particular, the confidence level of the participants tends to remain the same after the interaction (Figure 3.4B).

Finally, when the distance between opinions is very large (i.e., $\Delta O_{ij} > 1.1$), the strength of social influence diminishes progressively (Yaniv, 2004). In this zone, people seem to pay little attention to the judgment of another, presumably assuming that it may be an erroneous answer. Nevertheless, the others opinion is not entirely ignored, as the majority of people still choose the *"compromise"* heuristic when the partner is markedly more confident (i.e. $\Delta C_{ij} \leq -2$). Moreover, people who are initially very confident (i.e. $C_i \geq 5$) presumably begin to doubt the accuracy of their judgment and exhibit a high likelihood (of almost 70 %) of reducing their confidence level. Even more remote opinions are likely to be ignored entirely, but as this situation rarely occurs our data does not warrant a reliable conclusion here.

3.3 Simulations

3.3.1 The model

Taking these empirical regularities into account, we now elaborate an individual-based model of opinion adaptation and explore the collective dynamics of opinion change when many people influence each other repeatedly. To this end, we first describe the above influence map by means of a simplified diagram showing the heuristics that are used by most individuals according to ΔO_{ij} and ΔC_{ij} (Figure 3.3B). Alternatively, the same diagram can be characterized as a decision tree (Figure 3.3C). The model is defined as follows:

First, an individual notes the distance ΔO_{ij} between his or her own and a partners opinion and classifies it as near, far, or at an intermediate distance. For this, we used two threshold values of $\tau_1 = 0.3$ and $\tau_2 = 1.1$, assuming that the feedback is near when $\Delta O_{ij} < \tau_1$, far when $\Delta O_{ij} > \tau_2$, and at an intermediate distance otherwise. The numerical values of τ_1 and τ_2 were determined empirically from the influence map. Second, the individual considers the difference in confidence values ΔC_{ij} to choose among the three heuristics. Again, we define two threshold values α_1 and α_2 and assume that the individual decides to *"keep own opinion"* if $\Delta C_{ij} \geq \alpha_1$, to *"adopt other opinion"* if $\Delta C_{ij} \leq \alpha_2$, and to *"make a compromise"* otherwise. The three strategies can be formally defined as $R_i = O_i + \omega(O_j - O_i)$, where the parameter delineates the strength of social influence. Therefore, we have $\omega = 0$ when the individual decides to *"keep own opinion"*, and $\omega = 1$ when the individual decides to *"adopt"*. When the individual chooses the *"compromise"* strategy, that is when $0 > \omega > 1$, the average weight value $\bar{\omega}$ as measured from our data equals to $\bar{\omega} = 0.4$ (SD=0.24) , indicating that people did not move exactly between their initial estimate and the feedback (which would correspond to a weight value of 0.5), but exhibited a bias toward their own initial opinion (Soll and Larrick, 2009). Over all our data points, 53 % correspond to the first strategy ($\omega = 0$), 43 % to the second ($0 > \omega > 1$), and 4 % to the third ($\omega = 1$). The values of α_1 and α_2 depend on the distance zone defined before:

- When ΔO_{ij} is small, the others opinion constitutes a *confirmation* of the initial opinion. According to our observations, $\alpha = -5$ and $\alpha = -6$. Additionally, the confidence level C_i is increased by one point if $\Delta C_{ij} \leq -4$. As indicated by Figure 3.4A, C_i is also increased by one point with a probability $p = 0.5$ when $-4 \leq \Delta C_{ij} \leq 0$, and remains the same otherwise.

- When ΔO_{ij} is intermediate, the feedback has a significant influence on the subjects opinion. In this case, we set $\alpha_1 = 0$ and $\alpha_2 = -3$. The data shows that the confidence level is changed only if $\Delta C_{ij} \leq -3$ (Figure 3.4B). In this case, C_i increases with probability $p = 0.5$, and remains the same otherwise.

- When ΔO_{ij} is large, the thresholds are set to $\alpha_1 = -2$ and $\alpha_2 = -6$. This time, the confidence level decreases by one point when $\Delta C_{ij} \geq 4$, and remains the same otherwise.

Here, all the parameter values were directly extracted from the observations (Figure 3.3B and Figure 3.4).

3.3.2 Collective dynamics results

Having characterized the effects of social influence at the individual level, we now scale up to the collective level and study how *repeated* influences among *many* people play out at the population scale. Because the macroscopic features of the system are only visible when a large number of people interact many times, it would be extremely difficult to investigate this under laboratory conditions. Therefore, we conducted a series of numerical simulations of the above model to investigate the collective dynamics of the system.

The initial conditions of our simulations correspond to the exact starting configurations observed in our experiments (i.e., the precise opinion and confidence values of all 52 participants observed in the first experiment; also see Valori et al., 2012). In each simulation round, the 52 individuals are randomly grouped into pairs, and both individuals in a pair update their opinions according to the opinion of the other person, as predicted by our model. Thus, each individual is both a source and the target of social influence. We performed $N = 300$ rounds of simulated interactions, where N has been chosen large enough to ensure that the system has reached a stationary state. Here, we make the assumption that the decision tree that has been extracted from our experiment remains the same over repeated interactions. This assumption is reasonable to the extent that the outcome of the decision tree (i.e. the strategy that is chosen) depends on the confidence level of the individual, which is expected to change as people receive new feedback. In such a way, the strategies that will be selected by individuals are connected to the individual history of past interactions.

Figure 3.5 shows the dynamics observed for three representative examples of simulations. Although a certain level of opinion fragmentation still remains, a majority of individuals converge toward a similar opinion. As shown by the arrow maps in Figure 3.5, the first rounds of the simulation exhibit important movements of opinions among low-confidence individuals (as indicated by the large horizontal arrows for confidence lower than 3), without increase of confidence (as shown in Figure 3.8). After a certain number of rounds, however, a tipping point occurs at which a critical proportion of people meet up in the same region

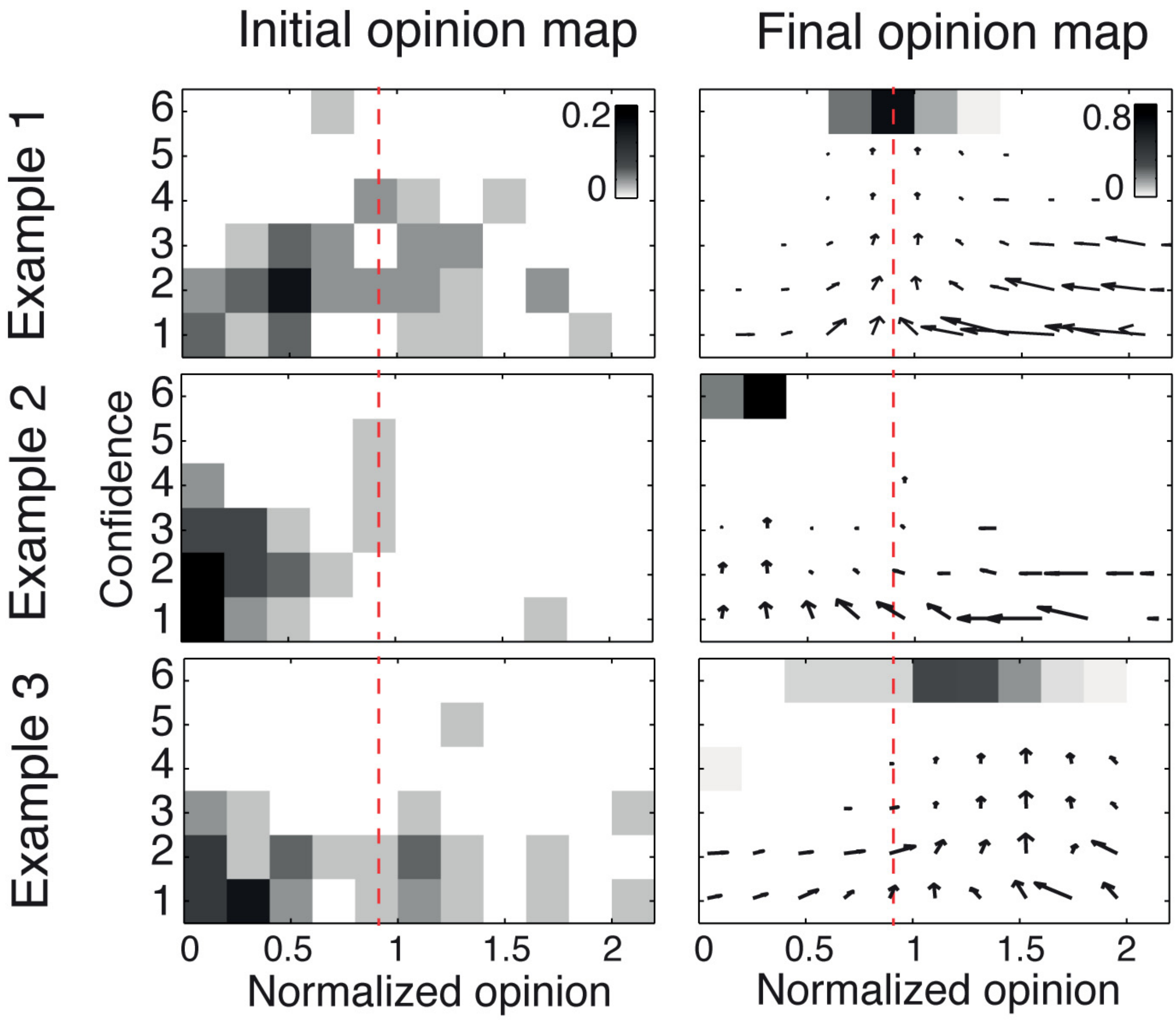


Figure 3.5: Three representative examples of the collective dynamics observed in the computer simulations. For each example, the initial opinion map is shown on the left-hand side (experimental data), and the final opinion map after N=300 rounds of simulations on the right-hand side. The opinion maps represent the proportion of individuals with a given opinion (x-axis) and a given confidence level (y-axis). As in Figure 3.1, the normalized opinion is the actual opinion divided by the true value. The correct answer is represented by the red dashed lines (corresponding to a value of 1). Outliers with normalized opinion greater than 2 are not shown. The arrow maps represent the average movements over both opinion and confidence dimensions during simulations. Examples 1, 2, and 3 correspond to the questions What is the length of the river Oder in kilometers? , How many inhabitants has the East Frisian island Wangerooge?, and How many gold medals were awarded during the Olympics in China in 2008?, respectively. The final convergence point may be determined by a dense cluster of low confidence individuals, as illustrated by Example 2 (majority effect), or by a few very confident individuals as in Example 3 (expert effect).

of the opinion space. This creates a subsequent increase of confidence in this zone, which in turn becomes even more attractive to others. This results in a positive reinforcement loop, leading to a stationary state in which the majority of people end up sharing a similar opinion. This amplification process is also marked by a sharp transition of the systems global confidence level (Figure 3.8), which is a typical signature of phase transitions in complex systems (Camazine et al., 2001).

An intriguing finding of our simulations is that the collective opinion does not converge toward the average value of initial opinions (a correlation test yields a nonsignificant effect with a coefficient $c = -.05$). The correlation between the convergence point and the median value of the initial opinions is significant ($p = .03$) but the relatively moderate correlation coefficient $c = 0.46$ suggests that this relation remains weak. Likewise, the system does not systematically converge toward or away from the true value (nonsignificant effect with a coefficient $c = .11$). Instead, the simulations exhibit complex collective dynamics in which the combined effect of various elements can drive the group in one direction or another. In agreement with previous works (Salganik et al., 2006), the collective outcome appears to be poorly predictable and strongly dependent on the initial conditions (Schelling, 1978). Nevertheless, we identified two major *attractors of opinions* that exert an important social influence over the group:

1. The first attractor is the presence of a critical mass of uncertain individuals who happen to share a similar opinion. In fact, when such a cluster of individuals is initially present in the system — either by chance or because individuals share a common bias — the rest of the crowd tends to converge toward it, as illustrated by Figure 3.5-Example2. This *majority effect* is typical of conformity experiments that have been conducted in the past (Asch, 1955), where a large number of people sharing the same opinion have a strong social influence on others.

2. The second attractor is the presence of one or a few highly confident individuals, as illustrated by Figure 3.5-Example3. The origin of this *expert effect* is twofold: First, very confident individuals exert strong persuasive power, as shown by the influence map. Second, unconfident people tend to increase their own confidence after interacting with a very confident person, creating a basin of attraction around that persons opinion (Couzin et al., 2005; Dyer et al., 2009).

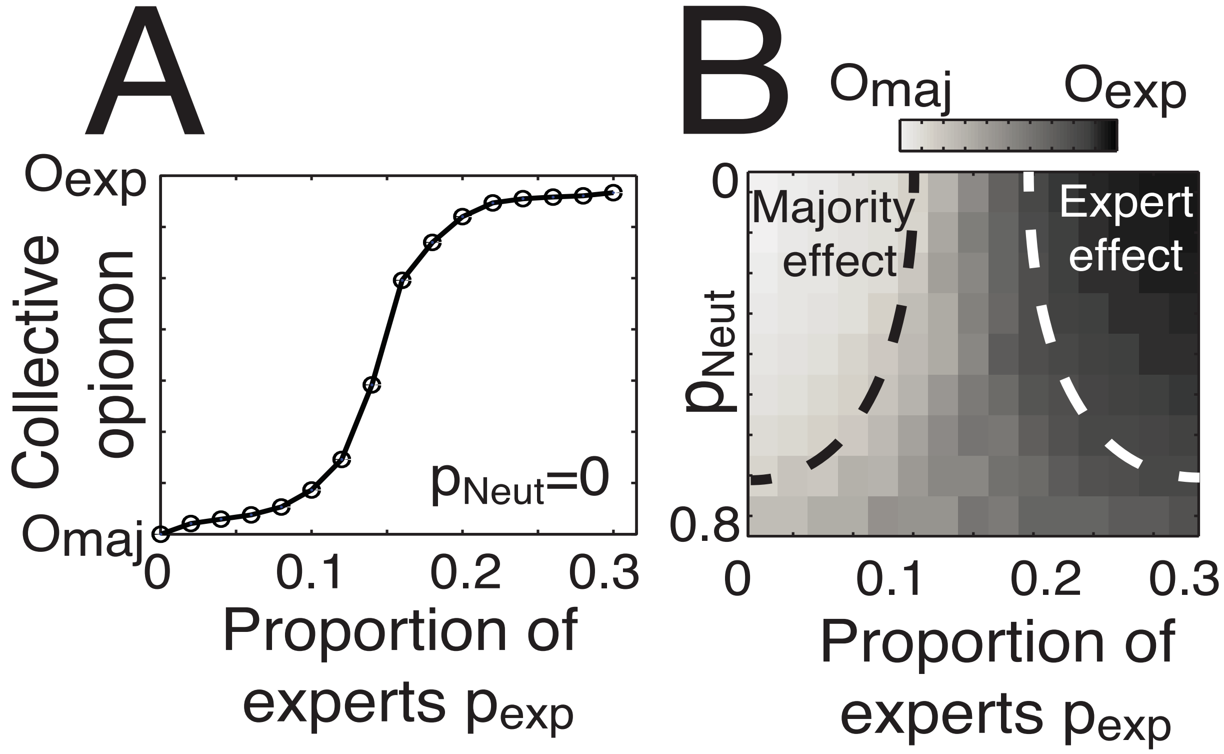


Figure 3.6: Which attractor dominates when the majority effect and the expert effect apply simultaneously? (A) The evolution of collective opinion when varying the relative proportion of experts p_{exp}, holding an opinion O_{exp} and a high confidence level $C_{exp=6}$, and the proportion of people in the majority group p_{maj} holding an opinion O_{maj} and a low confidence level randomly chosen in the interval C_{maj}=[1 3]. Here, the number of neutral individuals is fixed to $p_{Neut=0}$. (B) Phase diagram showing the parameter space where the majority or the expert effects applies, when increasing the proportion of neutral individuals p_{Neut} holding a random opinion and a low confidence level randomly chosen in the interval C_{uni}=[1 3]. The schematic regions delimited by black or white dashed lines show the zones where the collective opinion converges toward the majority or the expert opinion, respectively. In the transition zone, the collective opinion converges somewhere between O_{exp} and O_{maj}. In some rare cases, the crowd splits into two groups or more.

Our simulations show that the majority effect and the expert effect are not systematically beneficial to the group, as both attractors could possibly drive the group away from the truth (Figure 3.5-Example 2). What happens in the case of conflicting interests, when the expert and the majority effects apply simultaneously and disagree with each other (Figure 3.5-Example 3)? To investigate this issue, we conducted another series of simulations in which a cluster of low-confidence individuals sharing the same opinion O_{maj}, is facing a minority of high-confidence experts holding another opinion O_{exp}. As shown by Figure 3.6A, the majority effect overcomes the expert effect when the proportion of experts p_{Exp} is lower than a certain threshold value located around 10 %. However, as p_{Exp} increases from 10

%, to 20 % a transition occurs and the convergence point shifts from the majority to the experts opinion. Remarkably, this transition point remains stable even when a proportion p_{Neut} of neutral individuals (defined as people with random opinions and a low confidence level) are present in the system (Figure 3.6B). As p_{Neut} increases above 70 %, however, noise gradually starts to dominate, leading the expert and the majority effects to vanish. The tipping point occurring at a proportion of around 15 % of experts appears to be a robust prediction, not only because it resists to a large amount of system noise (Figure 3.6B), but also because a previous theoretical study using a completely different approach also reached a similar conclusion (Xie et al., 2011).

3.4 Discussion

In this work, we have provided experimental measurements and quantitative descriptions of the effects of social influence — a key element in the formation of public opinions. Our approach consisted of three steps: using controlled experiments to measure the effects of social influence at the scale of the individual, deriving a simple process model of opinion adaptation, and scaling up from individual behavior to collective dynamics by means of computer simulations.

The first result of our experiment is that participants exhibited a significant bias toward their own initial opinion rather than equally weighting all social information they were exposed to (Yaniv, 2004; Soll and Larrick, 2009). This bias is visible from the influence map shown in Figure 3.3, where the blue color corresponding to *"keep initial opinion"* is dominant and the red one corresponding to *"adopt the other opinion"* is rare. As shown in Figure 3.3B, the same trend has been transferred to the model. Moreover, even when the *"compromise"* strategy is chosen, individuals still give a stronger weight to their own initial opinion, which has also been implemented in the model. Therefore, contradictory feedback is typically underestimated — if not completely ignored — but opinions corroborating ones initial opinion trigger an increase in confidence. This observation is consistent with the so-called confirmation bias, namely, the tendency of people to pay more attention to information confirming their initial beliefs than information they disagree with (Nickerson, 1998; Baron, 2000). This result is also in line with early experiments showing that opinions tend to get reinforced by group discussions that involve people who initially share a similar judgment (Myers and Bishop, 1970). Likewise, the fact that individuals holding completely different beliefs exert very little influence on each other is consistent with the idea of "bounded confidence" — a modeling concept suggesting that social influence is negligible when opinions are initially too distant (Hegselmann and Krause, 2002; Lorenz, 2007). The presence of these elements confirms that our experimental design has indeed captured the fundamental mechanisms of social influence, and that factual questions can be used, to some extent, to study the fundamental features of opinion dynamics (Lorenz et al., 2011). In the future, an important challenge will be to evaluate how the influence map is shaped when emotions and subjective beliefs are more relevant (e.g. by using items about political opinions or beliefs that elicit strong convictions or emotions). Besides, another important follow-up study that should be conducted in the near future is the verification of our assumption that the decision tree observed at the first round of interaction remains identical over repeated interactions.

Scaling up from individual to collective behavior was achieved by means of computer simulations in line with existing approaches in the field of self-organization and complex systems (Schelling, 1978; Camazine et al., 2001; Castellano et al., 2009). Our simulations allowed us to unravel the precise mechanisms of opinion dynamics in large groups of people, which would have been practically impossible to characterize under laboratory conditions. In particular, an important ingredient underlying the collective dynamics but lacking in previous modeling approaches is the specific interplay between opinion changes and confidence changes. First, confidence serves as a sort of system memory. In fact, over simulation rounds, individuals are less easily influenced by others because their confidence level gradually increases as they receive new feedback. Therefore, simulated individuals do not constantly change their opinion but progressively converge toward a stable value in a realistic manner. Second, the increase of confidence supports the emergence of basins of attraction during collective opinion dynamics by boosting the attractive power of individuals sharing a similar opinion (Lorenz et al., 2011). This process often turns out to be detrimental to the group, because the local amount of confidence may grow *artificially* in a given region of the opinion space, which provides false cues to others and triggers a snowball effect that may drive the group in an erroneous direction. Interestingly, judgments of high confidence are good indicators of accuracy before social influence occurs, but no longer after people have been exposed to the opinion of others. It is remarkable that even a mild influence has a significant impact on the reliability of high confidence cues, as shown in Figure 3.2B. The main problem induced by social influence is that people tend to become more confident after noticing that other people have similar opinions. Therefore, high confidence is an indicator of accuracy when judgments are independent but becomes an indicator of *consensus* when social influence takes place (Koriat, 2012; Hertwig, 2012).

Our simulation results also identified two elements that can cause such amplification loops: the expert effect — induced by the presence of a highly confident individual, and the majority effect — induced by a critical mass of low-confidence individuals sharing similar opinions. Moreover, the presence of a significant number of neutral individuals holding a random opinion and a low confidence level around these two attractive forces tends to increase the unpredictability of the final outcome (Salganik et al., 2006). Therefore, neutral individuals make the crowd less vulnerable to the influence of opinion attractors, and thus less predictable. By contrast, recent studies on animal groups have shown that the presence of uninformed individuals in fish schools acts in favor of the numerical majority, at the expense of very opinionated individuals (Couzin et al., 2011).

Our simulations constitute a valuable tool that allows (i) unravelling the underlying mechanisms of the system, (ii) forecasting future trends of opinion change, and (iii) driving further experimental research and data collection. Nevertheless, it is important to note that the outcome of our simulations requires empirical validation in the future. This could be addressed, for instance, by means of empirical observations over the Web, where one would measure peoples opinion about a social issue over blogs and discussion forums and evaluate how the collective opinion changes over time (Wu and Huberman, 2004; Lazer et al., 2009). Alternatively, an online experimental approach such as the one elaborated by Salganik et al. seems well suited to the study of opinion dynamics under controlled conditions (Salganik et al., 2006). By quantifying the balance of power between the expert effect, the majority effect, and neutral individuals, our research can inform applications regarding the management of situations in which a small opinionated minority challenges a large population of uninformed individuals. For example, the model could help doctors convince a population of laypeople to adopt certain disease prevention methods or reversely prevent extremist groups from taking control of a large group of people.

3.5 References

Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, 193:33–35.

Baron, J. (2000). *Thinking and deciding*. Cambridge University Press.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100:992–1026.

Camazine, S., Deneubourg, J.-L., Franks, N., Sneyd, J., Bonabeau, E., and Theraulaz, G. (2001). *Self-Organization in Biological Systems*. Princeton University Press.

Castellano, C., Fortunato, S., and Loreto, V. (2009). Statistical physics of social dynamics. *Reviews of modern physics*, 81(2):591.

Couzin, I. D., Ioannou, C. C., Demirel, G., Gross, T., Torney, C. J., Hartnett, A., Conradt, L., Levin, S. A., and Leonard, N. E. (2011). Uninformed individuals promote democratic consensus in animal groups. *Science*, 334:1578–1580.

Couzin, I. D., Krause, J., Franks, N., and Levin, S. A. (2005). Effective leadership and decision-making in animal groups on the move. *Nature*, 433:513–516.

Deffuant, G., Neau, D., Amblard, F., and Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3:87–98.

Dyer, J., Johansson, A., Helbing, D., Couzin, I., and Krause, J. (2009). Leadership, consensus decision making and collective behaviour in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364:781–789.

Funk, S., Gilad, E., Watkins, C., and Jansen, V. (2009). The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences*, 106:6872–6877.

Funk, S., Salathé, M., and Jansen, V. (2010). Modelling the influence of human behaviour on the spread of infectious diseases: a review. *Journal of the Royal Society Interface*, 7:1247–1256.

Galam, S. (1997). Rational group decision making: A random field ising model at t = 0. *Physica A*, 238:66–80.

Galam, S. (2002). Minority opinion spreading in random geometry. *The European Physical Journal B*, 25:403–406.

Galam, S. and Jacobs, F. (2007). The role of inflexible minorities in the breaking of democratic opinion dynamics. *Physica A*, 381:366–376.

Galam, S. and Moscovici, S. (1991). Towards a theory of collective phenomena: Consensus and attitude changes in groups. *European Journal of Social Psychology*, 21:49–74.

Gigerenzer, G., Todd, P. M., and the ABC Research Group (1999). *Simple heuristics that make us smart*. Oxford University Press New York.

Hegselmann, R. and Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5(2).

Hertwig, R. (2012). Tapping into the wisdom of the crowdwith confidence. *Science*, 336:303–304.

Isenberg, D. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, 50:1141.

Kasperson, R., Renn, O., Slovic, P., Brown, H., Emel, J., Goble, R., Kasperson, J., and Ratick, S. (1988). The social amplification of risk: A conceptual framework. *Risk analysis*, 8:177–187.

Koriat, A. (2012). When are two heads better than one and why? *Science*, 336:360–362.

Latane, B. (1981). The psychology of social impact. *American psychologist*, 36:343–356.

Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., et al. (2009). Life in the network: the coming age of computational social science. *Science*, 323:721–723.

Liggett, T. M. (1987). *Interacting Particle Systems-An Introduction*. Springer.

Lorenz, J. (2007). Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics C*, 18:1819–1838.

Lorenz, J., Rauhut, H., Schweitzer, F., and Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, 108:9020–9025.

Martins, A. C. and Galam, S. (2013). Building up of individual inflexibility in opinion dynamics. *Physical Review E*, 87:042807.

Mäs, M., Flache, A., and Helbing, D. (2010). Individualization as driving force of clustering phenomena in humans. *PLoS computational biology*, 6:e1000959.

Mason, W. A., Conrey, F. R., and Smith, E. R. (2007). Situating social influence processes: Dynamic, multidirectional flows of influence within social networks. *Personality and social psychology review*, 11(3):279–300.

Moussaid, M., Garnier, S., Theraulaz, G., and Helbing, D. (2009). Collective information processing and pattern formation in swarms, flocks, and crowds. *Topics in Cognitive Science*, 1:469–497.

Myers, D. G. and Bishop, G. D. (1970). Discussion effects on racial attitudes. *Science*, 169:778–779.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2:175–220.

Salganik, M. J., Dodds, P. S., and Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311:854–856.

Schelling, T. (1978). Micromotives and macrobehavior.

Schweitzer, F. and Hołyst, J. (2000). Modelling collective opinion formation by means of active brownian particles. *The European Physical Journal B*, 15:723–732.

Slovic, P. (1987). Perception of risk. *Science*, 236:280–285.

Sobkowicz, P. (2009). Modelling opinion formation with physics tools: Call for closer link with reality. *Journal of Artificial Societies & Social Simulation*, 12(1).

Soll, J. and Larrick, R. (2009). Strategies for revising judgment: How (and how well) people use others opinions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35:780–805.

Sznajd-Weron, K. and Sznajd, J. (2000). Opinion evolution in closed community. *International Journal of Modern Physics C*, 11(06):1157–1165.

Valori, L., Picciolo, F., Allansdottir, A., and Garlaschelli, D. (2012). Reconciling long-term cultural diversity and short-term collective social behavior. *Proceedings of the National Academy of Sciences*, 109:1068–1073.

Wu, F. and Huberman, B. A. (2004). Social structure and opinion formation. *Computational Economics*, page 0407002.

Wu, F. and Huberman, B. A. (2007). Novelty and collective attention. *Proceedings of the National Academy of Sciences*, 104:17599–17601.

Xie, J., Sreenivasan, S., Korniss, G., Zhang, W., Lim, C., and Szymanski, B. (2011). Social consensus through the influence of committed minorities. *Physical Review E*, 84:011130.

Yaniv, I. (2004). Receiving other peoples advice: Influence and benefit. *Organizational Behavior and Human Decision Processes*, 93(1):1–13.

3.6 Appendix: Supplementary material

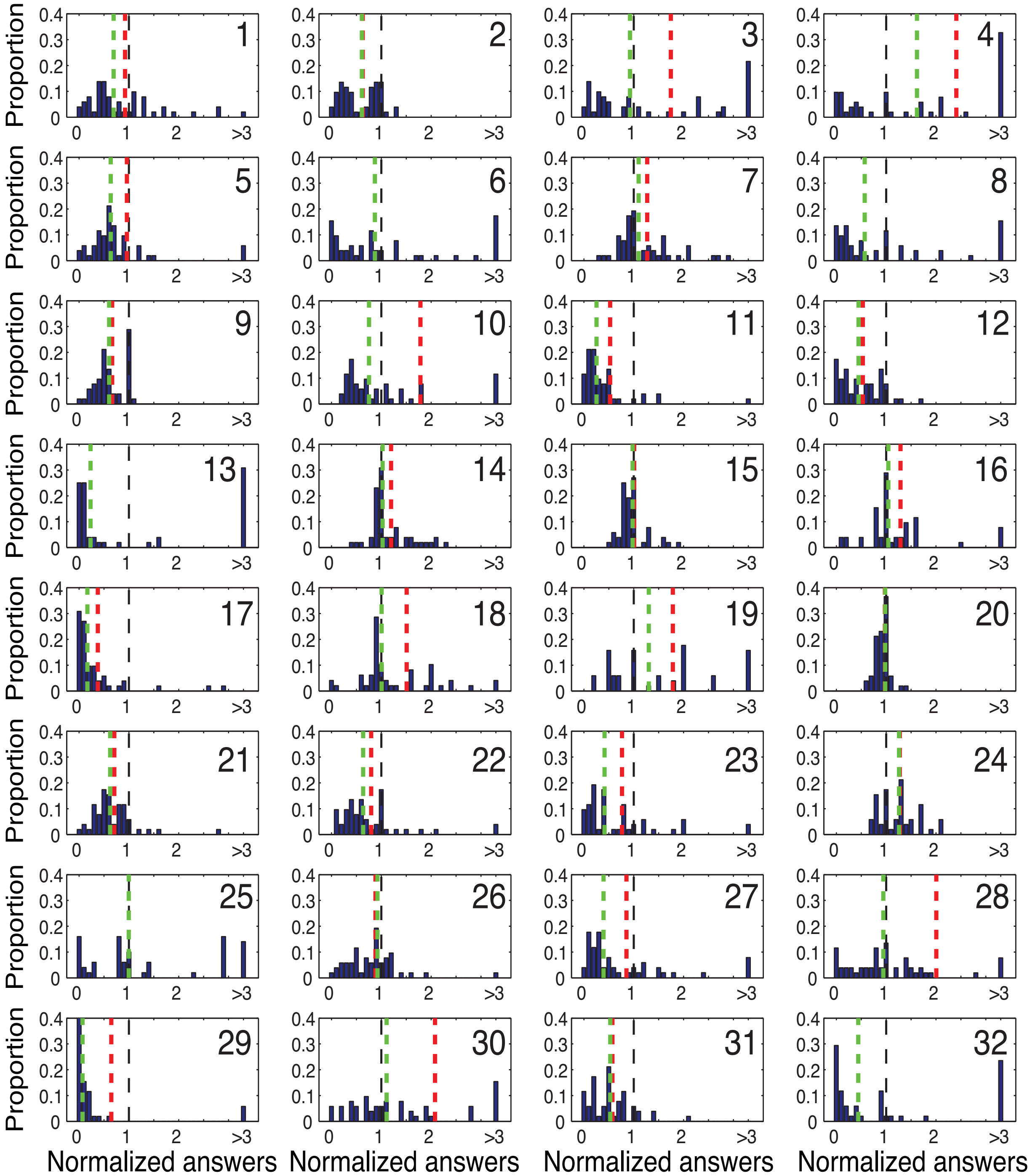


Figure 3.7: The distribution of answers for all 32 questions used in the first experiment (Experiment1, see Materials and Methods). The numbers on the upper right corner correspond to the question id, as indicated in the list of questions provided in the table in the Appendix 2. Question $id = 27$ has been used for illustrative purpose in the main text (Figure 3.1A). The normalized answer is the estimate of the participants divided by the true value. The black dashed lines indicate the correct answer (normalized value = 1). The red and green dashed lines indicate the mean and the median values of the distribution, respectively. The mean values lying farther than 3 are not indicated.

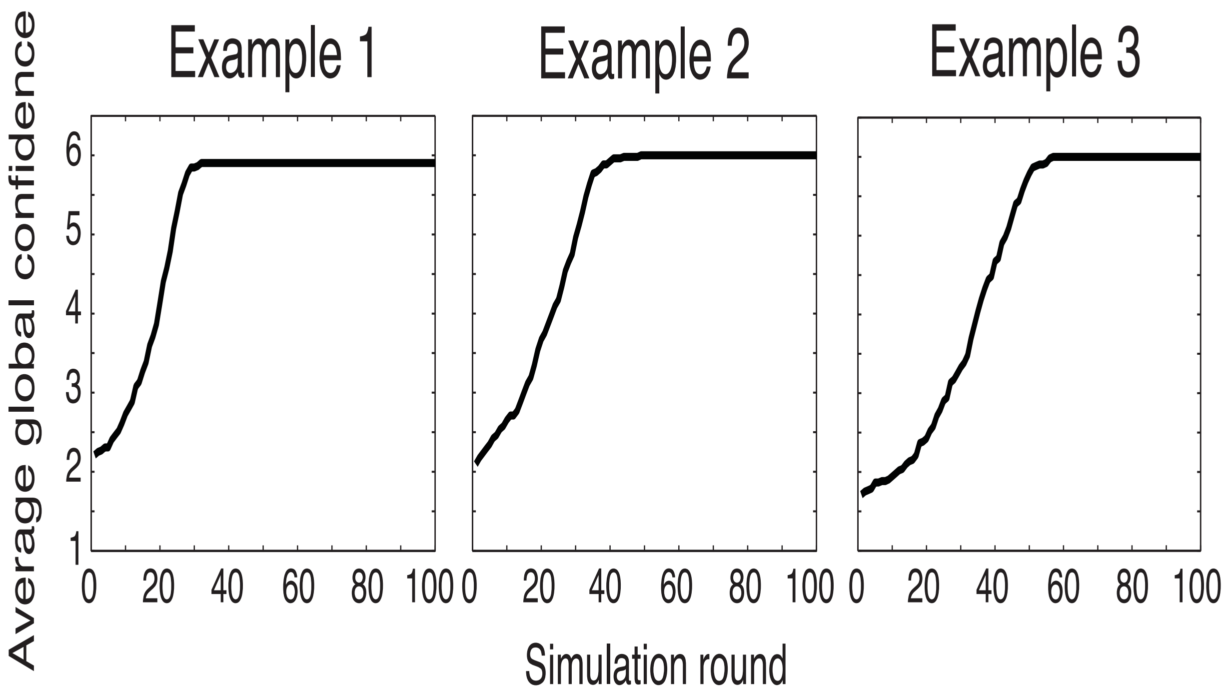


Figure 3.8: Three representative examples showing the evolution of participants confidence over simulation rounds. Examples 1, 2 and 3 correspond to those shown in Figure 3.4 in the main text. The average global confidence is computed by taking the mean value of confidence for all 52 participants. After a few rounds of simulation, a sharp transition occurs toward high confidence levels, attesting for the opinion amplification process.

3.7 Appendix 2: Materials

Id	Question	Answer
1	What is the length of the river Oder in kilometers?	866
2	What is the height of the Fernsehturm in Berlin?	368
3	What is the height of Uluru (Ayers Rock) in Australia (in meters, ASL)?	863
4	How deep is the Baltic Sea at its deepest point (in meters)?	459
5	How long is the boarder between Germany and Switzerland (in km)?	316
6	What is the population density in inhabitants per square kilometer(2009)?	229
7	What is the distance between Berlin and London? (in kilometers)?	910
8	How many inhabitants has the East Frisian island Wangerooge (retrieved 2010)?	919
9	How many members has the German Bundestag according to the law (without overhang seats; in 2012)?	598
10	How many universities are there in Germany (in 2010/2011)	106
11	How many active nuclear power stations are there in Europe (retrieved 2011)?	196
12	How many countries take part in the general assembly of the United Nations as active members?	193
13	How many assaults per 100.000 inhabitants were officially registered in Germany in 2007?	618
14	What is the monthly amount of basic security benefits for full-aged, single job seekers (Hartz IV, in 2012)?	364
15	What is the monthly amount of child benefit in Germany for the first child (in 2012)?	184
16	How much does an iPad2 with 16GB cost (RRP)?	479
17	How many gold medals where awarded during the Olympics in China in 2008?	302
18	What is the world record in high jump for men (in centimeters)?	245

19	How long was the track for running events in a stadium in the Ancient Olympic Games (in meters)?	192
20	What is the maximum speed ever reached during a Formula One race (Grand Prix; in km/h)?	370
21	What is the world record in ski jumping of men (in meters)?	247
22	How long is the longest tee in golf so far (in meters)?	471
23	How many sportmen have taken part in the first modern Olympic Games in Athens in 1896?	241
24	How many kilograms has a sumo wrestler to weight at least for being admitted as heavyweight?	115
25	What is the speed of sound in the air (MSL)?	343
26	How many bones does an adult human have?	206
27	What is the melting temperature of aluminium (in degrees Celsius)?	660
28	How many degrees Fahrenheit are 100 degrees Celsius?	212
29	How many earthquakes with value more than 6 on the Richter scale happen in an average year worldwide?	150
30	How many calories does a liter of Coca Cola contain (kcal)?	420
31	How many earth(days) has a year in the Mars?	687
32	How many times larger is the diameter of the sun compared to the diameter of the earth?	109

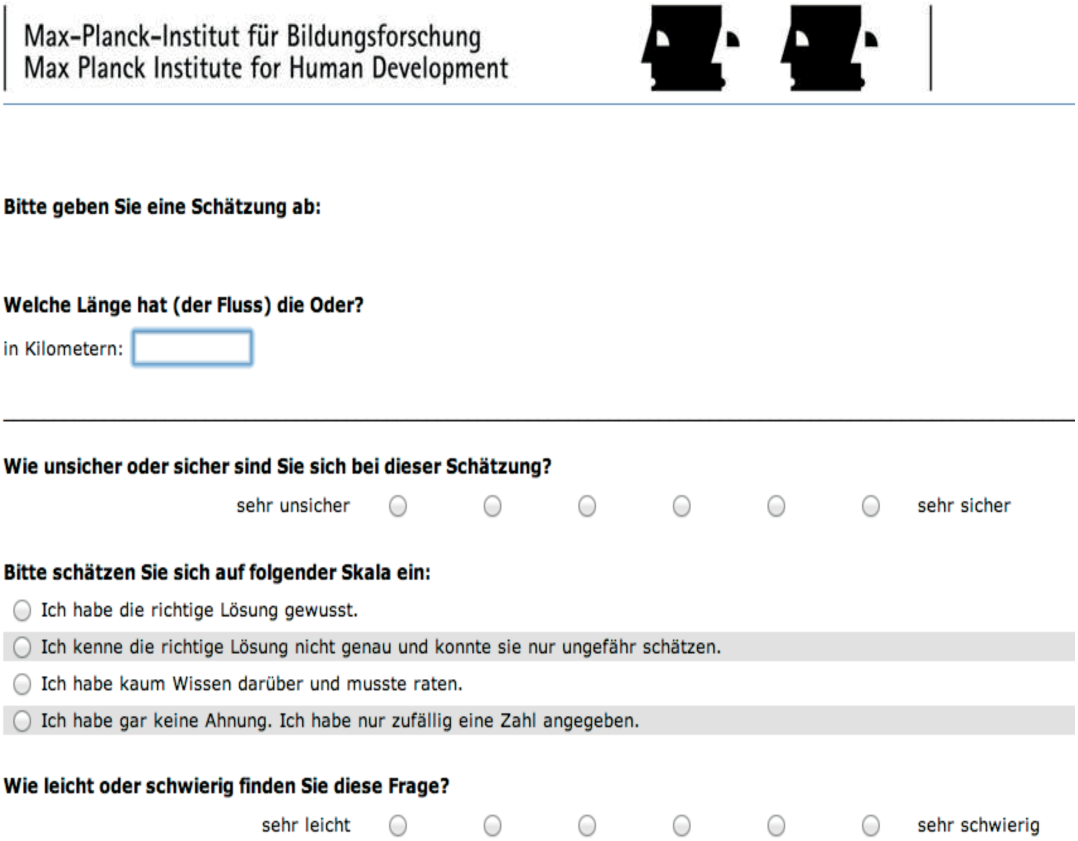
Table 3.1: The 32 general knowledge questions used in our experiments.

Instructions-Experiment 1

Welcome! In the study Quantitative estimates we investigate human decision-making. In what follows you will be shown 32 estimation tasks and you are asked to provide your quantitative estimates.We would like to ask to answer all the estimation tasks spontaneously and truthfully without any external help (e.g lexica or the internet).
You will not be evaluated according to the correctness of your estimates. As a result, if you use other sources of information, you will only harm the quality of the results of this study. The correct answers will be presented to you at the end of the experiment.
All the answers will be evaluated anonymously and without relation to your personal information. The study will last approximately 40 minutes. You will receive 8 Euros for your participation. In case you terminate the experiment too early we will not compensate you.

Instructions-Experiment 2

Welcome! In the study Quantitative estimates we investigate human decision-making. In what follows you will be shown 30 estimation tasks and you are asked to provide your quantitative estimates. We would like to ask to answer all the estimation tasks spontaneously and truthfully without any external help (e.g lexica or the internet).
Each task has two parts. In the first part you will be asked to provide a first spontaneous estimate. In the second part you will receive information about the estimates of another participant or a group of participants on the same task and you will be asked once more to provide your estimates.
You will not be evaluated according to the correctness of your estimates. As a result, if you use other sources of information, you will only harm the quality of the results of this study. The correct answers will be presented to you at the end of the experiment.
All the answers will be evaluated anonymously and without relation to your personal information. The study will last approximately 40 minutes. You will receive 8 Euros for your participation. In case you terminate the experiment too early we will not compensate you.



Max-Planck-Institut für Bildungsforschung
Max Planck Institute for Human Development

Bitte geben Sie eine Schätzung ab:

Welche Länge hat (der Fluss) die Oder?

in Kilometern:

Wie unsicher oder sicher sind Sie sich bei dieser Schätzung?

sehr unsicher ○ ○ ○ ○ ○ ○ sehr sicher

Bitte schätzen Sie sich auf folgender Skala ein:

- ○ Ich habe die richtige Lösung gewusst.
- ○ Ich kenne die richtige Lösung nicht genau und konnte sie nur ungefähr schätzen.
- ○ Ich habe kaum Wissen darüber und musste raten.
- ○ Ich habe gar keine Ahnung. Ich habe nur zufällig eine Zahl angegeben.

Wie leicht oder schwierig finden Sie diese Frage?

sehr leicht ○ ○ ○ ○ ○ ○ sehr schwierig

Figure 3.9: Screenshot of the first experiment. The participants answered 32 general knowledge questions and reported their confidence and the experienced difficulty of the questions.

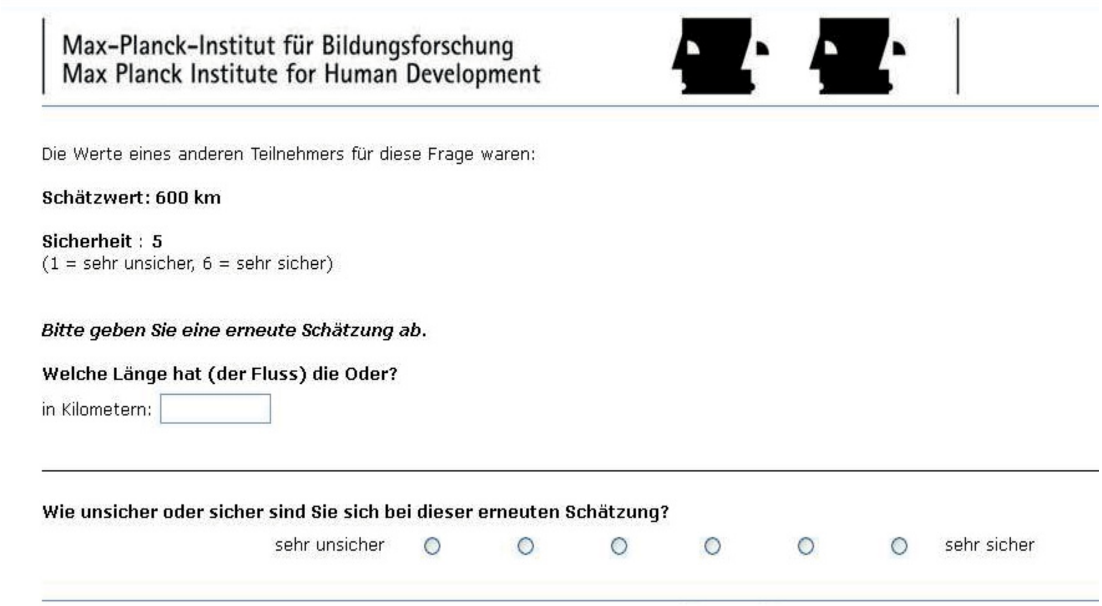


Figure 3.10: Screenshot of the second stage of the second experiment. After giving their first answer and reporting their confidence the participants were exposed to the answer and confidence level of another participant and were asked to revise their initial estimate.

Chapter 4

The collective dynamics of sequential search in markets for cultural products

Pantelis P. Analytis, Hrvoje Stojic and Mehdi Moussaïd

Chapter Abstract

Markets for cultural products are characterized by highly skewed popularity distributions. In addition, it is notoriously hard to predict the success of specific cultural products. We show that both these emerging properties of the markets can be captured in a parsimonious model of social interaction in which agents with diverse yet correlated preferences search the alternatives in order of popularity and choose the first alternative with utility higher than a certain satisficing threshold. The model goes beyond existing accounts of the market dynamics in that (i) it suggests a novel and psychological mechanism — sequential ordered search — by which social influence plays out (ii) and it allows us to study the welfare properties of the market in multi-alternative environments. In agent-based simulations we systematically varied the diversity of preferences in the agent population and the satisficing threshold employed by the agents and examined the implications for the market dynamics. We found that searching by popularity leads to an increase in inequality and unpredictability

in the market compared to cases in which agents have perfect knowledge of the alternatives or search among them at random. Further, we found that in some cases diversity of preferences in the agent population leads to an increase in the aggregate welfare enjoyed by the agents in the market.

4.1 Introduction

In markets for cultural products such as books, music recordings, movies or scientific articles, a great number of alternatives compete for consumer attention. Consumers tend to consider only a small subset of the products available in the market and eventually buy only a few of the alternatives from those they carefully considered. There are two well-documented empirical observations about these markets. First, they are characterized by highly skewed popularity distributions — a few "hit" products conquer a large market share, while many more attract only a few consumers (Merton, 1968; Mitzenmacher, 2004; Newman, 2005). Second, the success in the market is notoriously hard to predict before a product's release, regardless of the tools at the forecasters' disposal (De Vany, 2004; Goel et al., 2010). Well-financed blockbusters may eventually fail to find an audience, whereas music albums from unknown bands may become classics. In the past, different methodologies have been used to reproduce in a model these two properties of these markets (e.g. Adler, 1985; Chung and Cox, 1994).[1]

Recently, a set of groundbreaking experiments conducted by Salganik et al. (2006) indicated that social information in the form of other people's choices leads to an increase in inequality and unpredictability in the market. In addition, the experiments showed that the large consumption inequalities observed in the markets reflect actual differences in quality, but only partly so. Salganik et al.'s work demonstrated that previous modelling accounts of the markets for cultural products were at best incomplete. Further, the experiments hinted to the rich welfare dynamics of these markets. Clearly, hit alternatives that are eventually consumed by a large part of the population have a great impact on aggregate welfare. Consider scientific articles or whole research programs that accumulate citations and are followed by numerous graduate students, while work of higher quality perhaps remains undiscovered. To design more efficient market mechanisms, it is crucial to understand how the popularity dy-

[1]The authors would like to thank Perke Jakobs, Mirta Galesic, Russell Golman, Shenghua Luan and Thomas Stork for their remarks on earlier versions of the paper and the members of the ABC group for their feedback during internal presentations. The authors would also like to thank Anita Todd for editing the manuscript.

namics arise from individual behavior, through which some of the alternatives become hits while many more remain shrouded in oblivion. How to model the decision processes at the individual level identifying the role of social information, remains an open question.

The first attempts to model popularity dynamics were based on models of stochastic processes that fit market dynamics well, but did not take into account quality differences or postulate any decision-making processes (e.g. Price, 1976; Chung and Cox, 1994). Later, models that spelled out the individual decision-making process and social influences brought new conceptual insights to the field (e.g. Adler, 1985; Banerjee, 1992), but abstracted away many of the properties of the markets. More recent models, to a large extent inspired by Salganik et al.'s (2006) experiments, represent significant progress (Borghesi and Bouchaud, 2007; Hendricks et al., 2012; Krumme et al., 2012; Denrell and Le Mens, 2013). However, the modelers either focused on the specifics of Salganik et al.'s experiments or postulated strict assumptions to derive precise analytical solutions. This leaves a lot of room for improvement in terms of (i) the external validity and scope of the model (ii) drawing conclusions about welfare outcomes in the markets for cultural products.

We adopt and extend a well-established framework of individual decision-making in economics, marketing and psychology — sequential search (Simon, 1955a; DeGroot, 1970). The agents in our model use a simple psychological search rule: They consider the alternatives in the order of aggregate popularity information. This allowed us to fit important characteristics of cultural goods markets better than previous models did and improve the external validity. In our model, consumers (agents) sample the alternatives one after another until they encounter an alternative with quality higher than a satisficing threshold. We relied on agent-based simulations rather than deriving closed-form analytical solutions. This gave us the flexibility to explore more representative market scenarios. Scaled up to the market level the model accounts for the emerging properties of the markets for cultural products. By varying structural characteristics of the environment — the diversity of preferences in the agent population and the cost of sampling one alternative — we generated predictions for the inequality and unpredictability in the market that could be tested against data from the field and be used to guide future experimentation. We compared our model against random search, a well-studied choice model for which optimal solutions exist. For cost of search equal to zero the two processes are equivalent to the perfect information hypothesis. However, they strikingly diverge as the costs of search increase. We analyzed the welfare outcomes of the market and how they depend on parameters of the model. We found a surprising beneficial effect of moderate amounts of diversity of preferences on the welfare of market participants. Finally, since it is an agent-based model, the framework is flexible and other psychological mechanisms, such as product awareness or network externalities, can be explored in the future.

4.1.1 Literature review

A number of methodologies have been used to model the markets for cultural products. Some explicitly attempt to reproduce the macroscopic patterns observed in actual markets, while other take an abstract approach and attempt to provide general insights into individuals' decision-making processes and their welfare implications. In what follows we categorize the literature into three distinct groups. The first consists of models formulated as stochastic processes that with time produce highly skewed popularity distributions. The first model of this kind was presented by Yule (1925) to describe the process of speciation in biology. In the 1950s Simon (1955b) developed a general stochastic process model that could account for skewed distributions observed in many empirical phenomena in the social and natural sciences. Price (1976) presented the first model that explicitly attempted to capture the popularity dynamics of citations of academic publications — a cultural market par excellence. His model was expressed as a Polya's Urn, in which the discrete probability of an event increases every time that the event occurs. This leads to a positive feedback loop that over time makes some events very likely to occur while other events become very unlikely. Working along similar lines, Barabási and Albert (1999) developed a well-known model of preferential attachment. With proper parameterization, their model can account for the inequality in distribution of hyperlinks among websites on the Internet but more abstractly it can be employed to capture any type of popularity dynamics. The last two models capture the emerging dynamics of the market and the resulting popularity distributions quite accurately. However, they abstract away the quality differences of the alternatives as well as the decision-making processes of the individuals participating in the market. The alternatives in these models are described only in terms of the dynamically acquired popularity. Thus, although they provide a good summary of the properties of the market on the macro level, they are insufficient for identifying what aspects on the individual level give rise to the rich-get-richer dynamics and the market outcomes.

The second group consists of models in which the inequality and unpredictability in the market are driven by inherent or socially acquired differences in quality or utility. These models try to map the important aspects on the individual level that lead to observed market outcomes. Rosen (1981), for example, used a convex relationship between quality and market success to explain why small differences in product quality or utility can lead to large differences in market success. However, accounts based exclusively on quality differences cannot explain the difficulty of predicting success ex-ante in these markets. A breakthrough in explaining unpredictability came with incorporating social influences, such as positive network externalities, into the models: Leibenstein (1950), Adler (1985), Elster (1989) and Brock and Durlauf (2001) have shown that when agents directly derive utility from coordinating on the same choices as other agents, unequal and unpredictable consumption distributions can arise. For example, people might buy a popular book because they will enjoy it even more due to all the conversations they we will have about it with their friends and colleagues. Arthur (1989) has shown how a similar principle works on the production side — the more consumers adopt a certain product early on, even by chance alone, the more resources the company will have, and the product's quality will increase. This creates a positive feedback loop, which further increases the number of consumers. Network externalities certainly play a role in cultural goods, but they are unlikely to be the whole story. This is exactly what the recent experiments conducted by Salganik and his collaborators (Salganik et al., 2006; Salganik and Watts, 2008, 2009) have shown. They found that a different form of social influence plays a crucial role — social information in the form of other people' choices enters the agent's decision process and strongly influences inequality and unpredictability in the market.

Why do we give such weight to Salganik et al.'s (2006) experiments? They designed an artificial music market called Music Lab with thousands of participants. Even though participants knew the market was part of a science project, the experiment was very realistic and its design simulated an actual online market. Participants in the market decided sequentially to listen to and download any number of 48 songs from relatively unknown bands. In the baseline condition the participants chose to listen and download songs from a randomly ordered list. In one social influence condition Salganik et al. introduced social influence by presenting next to the song the number of times the song was downloaded by other participants. In another condition, they additionally ordered the songs according to their popularity, as in a bestseller list. They repeated the experiment with the same initial conditions several times with new samples of participants each time. This allowed them to measure the inequality and unpredictability in the emerging market. Given the strength of their experimental design and its ability to measure preferences without social influence, they have mounted a convincing case that inequality and unpredictability in the market increase substantially as a result of the social information available to the market participants.

Salganik et al. (2006) had clear experimental hypotheses, but did not focus on the decision-making processes of the participants and did not postulate a precise model that would explain their results. To fill this gap, Borghesi and Bouchaud (2007) and Krumme et al. (2012) developed models tailored to the design of the experiment. Borghesi and Bouchaud (2007) constructed a model inspired by the Ising model of ferromagnetism and employed it to replicate the end of simulation results of the Music Lab experiment. In their model, like in some of the models discussed in the second category, social influence can be seen as a network externality parameter, which is added to the available public and private information when the agents choose whether to download a song. By contrast, Krumme et al. (2012) described the decision to download a Music Lab song as a two-step process. The agents first choose whether to sample a song and then whether to download it. In their model, social influence enters as an availability parameter that controls for the probability that an item will be sampled given its position on the screen. The models reproduce the popularity dynamics observed in the Music Lab experiments, but (i) they require further specification in order to generate predictions for markets with different conditions than in the Music lab experiment and (ii) they do not consider the implications of social interaction for the welfare of the participants in the market. Krumme et al. provided evidence that social influence is informational. This leads us to a third group of models that like in the model presented in Krumme et al. highlight the role of social information in individual decision processes.

In contexts in which decision makers choose sequentially, Banerjee (1992) and Bikhchandani et al. (1992) have shown that the decisions of very few individuals in the beginning of the decision-making sequence determine whether an alternative will be selected or rejected. Hendricks et al. (2012) developed a model based on that insight that captures the main characteristics of real-world cultural markets. They populated the market with agents following a simple search model, where agents have heterogeneous preferences and observe only the aggregate purchase history. They found that availability of social information in this form affects the probability that herds will occur, depending on the quality of the products. The probability that low-quality products will become hits goes to zero, while for high-quality products the probability that a sub-optimal product becomes a hit is greater than zero. In our opinion this model comes the closest to representing the crucial aspects of cultural goods markets. However, the decision to search and then select an alternative is made individually for each option. In markets swarmed by a great number of alternatives this is very strong an assumption. In addition Hendricks et al. made several simplifying assumptions to derive analytical solutions. For example, alternatives in their models can be either low or high quality and agents know the decision processes of all other agents and update their decisions in a fully rational way. Imposing these constraints again impedes the external validity of the model and makes it difficult to generalize results to slightly different but equally interesting market scenarios. This also limits their analysis of welfare consequences of social influence in the market.

4.1.2 Our modeling approach

We sought to develop a simple model of an individual decision-making process with assumptions that reflect human cognitive capacities and the most salient characteristics of the markets for cultural products. We started with some key observations about the markets: People can learn the true utility of a product fairly quickly. For example, before a purchase consumers can listen to a song, watch a movie trailer, skim through a book, or read an abstract of a scientific article to assess its utility. However, to learn the utility consumers examine the products one by one and pay a non-negligible cost in terms of time or effort. For instance, people can read only one abstract at a time and usually pay a cost in time spent reading and not doing something else. Further, consumers are constrained by their time or budget to consume only a limited number of these products. Finally, simply because there are a great number of products in the environment, consumers do not have the time resources to examine each one before making a choice. Thus, consumers necessarily consider only a subset of the total number of alternatives available in the environment.[2] These elements of the decision making process can be neatly and formally captured by a sequential search model (e.g. DeGroot, 1970).[3] In our model each agent makes decisions following this process. Intuitively, such an agent stops searching after encountering a good-enough alternative, although better alternatives may remain undiscovered in the market. This can be expressed in the form of a stopping threshold: The agent continues to sample alternatives until it encounters one that satisfies this threshold. With certain assumptions satisfied and when agents make decisions in isolation, optimal solutions can be computed. Importantly,

[2]These observations lead us to believe that the simultaneous examination of alternatives used in some previous models (Borghesi and Bouchaud, 2007) is an oversimplification. Similarly, modeling the process as a choice of a single good (Hendricks et al., 2012) is also oversimplifying since decisions might depend on alternatives examined previously.

[3]An early model of this kind was put forward by Simon (1955a) as a boundedly rational alternative to omniscient utility-maximizing agents. Shortly thereafter, a mathematically rigorous theory of optimal sequential search behavior was developed in statistics (DeGroot, 1970). The models of sequential search have been widely applied — for example, exploring the search for jobs (McCall, 1970), lower prices (Telser, 1973) or new technologies (Muth, 1986).

there are experimental results showing that individual behavior in such problems can be approximated by threshold decision strategies (Rapoport and Tversky, 1966; ?; Hey, 1987; Lee, 2006).

Further, in markets for cultural products social influence shapes the interaction among consumers — in our highly interconnected world agents often possess some information about the popularity of the alternatives, or their information on available alternative is influenced by popularity. This is reflected in popularity of "best of" lists for movies, books, songs, albums and so on. Popularity affects the probability that a person will hear a song on the radio (Hendricks and Sorensen, 2009); people are more likely to talk about popular products with their friends; news websites are more likely to report on such products; and in online stores items are often by default ordered according to their popularity, or consumers can easily do so. At the cognitive level, popularity could be reflected in the fluency with which agents process the alternatives in memory (Hertwig et al., 2008).

There are several ways to implement information about popularity in the model. We took the simplest route by influencing the sampling process rather than directly the choices (Denrell and Le Mens, 2013); popularity determines the order in which the alternatives are searched.[4] As in a bestseller list, the agents start by examining the alternative that has been selected by most agents and move down to less popular alternatives. Importantly, this is not a "globally rational" method, where agents would have access to all the decisions of previous agents and their order of searches. Instead, the agents behave as naive statisticians, not questioning where their sample comes from (Fiedler and Juslin, 2006) but they still make rational decisions based on their sample (Denrell and Le Mens, 2007). This approach is in line with the observations from the Music Lab experiment (Krumme et al., 2012). We will show that a model based on a population of agents whose decision making is driven by such a simple model already produces promising results that capture the stylized facts of the markets.

[4] A basic assumption of sequential search models is that the decision-makers search the alternatives in a random fashion (DeGroot, 1970). In our simulations we used this model as a reference frame where agents are under no social influence.

Conceptually, our model incorporates many of the notions of existing social influence models such as sequentiality in choice, threshold rules, and heterogeneity in preferences (e.g. Granovetter and Soong, 1986; Arthur, 1989; Banerjee, 1992; Hendricks et al., 2012) and illustrates how these can be operationalized in multi-alternative choice environments. However, it goes beyond these models as it illustrates that a simple cognitive mechanism — limited sequential search — captures the emerging properties of the market such as skewed popularity distributions, outcome unpredictability and the imperfect relation between perceived quality and popularity better than previous theoretical accounts. Note that our model is the first that explicitly lays down two types of sequentiality. *Within agent sequentiality*, found in search models (e.g. DeGroot, 1970) and which refers to the agents looking at the alternatives one after another and stopping once they find a satisficing one and *between agent sequentiality* found in informational cascade and herding models (e.g. Banerjee, 1992), which refers to the agents deciding one after another. Previously, models of social influence have taken into account only the second.

The remainder of the paper is structured as follows: We present the model in Section 4.2, where we also provide a basic analysis and discuss the assumptions. In Section 4.3 we present the results of an agent-based simulation over a broad range of parameters of the model. In Section 4.4 we discuss the results, present some applications and extensions, and finish with concluding remarks.

4.2 The model

4.2.1 The environment

There are N alternatives (or goods) $X_1, ..., X_N$ in the market, populated by M agents $A_1, ..., A_M$. The alternatives have an objective utility component u_o, which is identical for all the agents, and a subjective component u_s, which is agent specific. The overall utility of an alternative u is then a simple sum of these two components $u = u_o + u_s$. The objective component u_o of each alternative is a draw from an independent and identically distributed (iid) random variable, normally distributed with mean μ and variance σ_o^2. The subjective component u_s of each alternative is iid normal with the same mean μ, but with a different variance σ_s^2. The overall utility is then a sum of two draws from iid normal variables, which itself is an iid normal variable with variance $\sigma^2 = \sigma_o^2 + \sigma_s^2$.

The agents encounter alternatives sequentially and they can learn the utility of an alternative u_n only by sampling it and paying a cost c. We refer to c throughout the text as *search cost*. The agents can sample as many alternatives as desired, but they can choose only one of the alternatives they have sampled. An example of sampling behavior could be listening to a couple of songs or skimming through several books and finally opting for one of them. We assume that the search cost is constant and that there is no post-sampling uncertainty — by sampling the alternative the agent learns its true utility. The framework can be easily extended to cases where the agents can choose more than one alternative or cases where some uncertainty remains even after sampling.

An agent's return function can be summarized as,

$$R_n(u_1, u_2, ..., u_n) = max(u_1, u_2, ...u_n) - nc \tag{4.1}$$

Note that the returns depend on the utility of the best alternative discovered so far and the search cost. The returns from sampling one more alternative can be formulated as,

$$\begin{aligned} R(u_n, c) &= Pr(u_n > max(u_1, u_2, ...u_{(n-1)})) \times \\ & E(u_n - max(u_1, u_2, ...u_{(n-1)}) | u_n > max(u_1, u_2, ...u_{(n-1)})) - c \end{aligned} \quad (4.2)$$

When viewed on the level of a single agent, this is a classical optimal stopping problem — examined extensively in statistics and economics (e.g. DeGroot, 1970) — in which after examining a new alternative the agents decide whether to sample further or to stop search. Optimal stopping problems with a payoff function defined this way are called stopping problems *with recall*.

In studying the effects of social influence on the behavior of agents, it is important to have baselines against which the effects of search behavior and social influence on the structure of the market and the welfare outcomes can be measured. In network theory, for example, the random graph model continues to serve the role of baseline against which other network structures are compared (e.g. Watts and Strogatz, 1998).[5] For this reason, we examine two types of search scenarios — one in which agents do not have access to popularity information and one in which they do. We call the first market a random search market (random search in short), and the second a popularity heuristic market (popularity heuristic in short). When the cost of search is 0 both these processes converge to the perfect knowledge assumption of neoclassical models (for a similar experimental comparison see Reutskaja et al., 2011).

4.2.2 Random search

The agents do not have access to any social information and they search the alternatives in random order. This scenario has been studied extensively and complete theoretical treat-

[5]This is straightforward to do in our model and simulations, but empirically it is a challenging problem. One of the important contributions of Music lab experiments by (Salganik et al., 2006) was exactly a successful measurement of preferences without social influence.

ments can be found in Chow et al. (1971). An optimal threshold T can be found for which the return from sampling one more alternative is zero, $R(u_n, c) = 0$. A standard result is that as the cost of sampling increases, the optimal threshold T decreases. Intuitively, for a high cost, an agent is willing to sample more only if the highest u discovered thus far came from the lower part of the distribution. Moreover, the return is always positive as long as the utility of the best sampled alternative is lower than the optimal threshold.

When sampling from a distribution with known variance and mean, the optimal threshold depends only on the cost of sampling and can be written as $T(c)$. This result holds irrespective of the total cost of search paid so far or the number of alternatives that remain to be sampled. The optimal sampling policy for agent i can be formally summarized as:

$$\begin{aligned} &\text{if} \quad u_{ni} > T_{opt}(c), \quad \text{stop sampling} \\ &\text{if} \quad u_{ni} \leq T_{opt}(c), \quad \text{continue sampling} \end{aligned} \tag{4.3}$$

The agents $A_1, ..., A_M$ make their choices sequentially. An agent is chosen randomly without replacement. The agent then searches through alternatives and chooses one of them. This process is repeated until all M agents have made their choice and the market outcome is known. Since agents do not make use of social information we use random search as a reference frame to evaluate social influence.

4.2.3 Social interaction

Here we describe how, exactly, social information is created in the market. This information will then be used by agents following the popularity heuristic.

After the first agent, a second or m^{th} agent is chosen to make their choice, again at random without replacement. Now the agent can observe the choices made by previous agents and potentially benefit from them. However, the agents cannot observe a detailed history of choices. Instead they observe the summary information of how many times each alternative was chosen.[6] In other words, there is a popularity vector $P = \{P_1, ..., P_N\}$ that records the choices of the corresponding alternatives $X_1, ..., X_N$ and it is updated whenever an agent makes a choice. Examples of such popularity information would be product sales, number of song downloads, or number of article citations. Thus, the only point of social interaction is via the publicly available information on popularity P.

4.2.4 Popularity heuristic

The agents sample the alternatives $X_1, X_2...X_n$ in decreasing order of popularity $P_1 > P_2 > ... > P_n$, where popularity is defined as the number of times that an alternative has been selected in the past. When there is a tie, the agents choose which one to sample next at random. Thus popularity heuristic can be seen as a single attribute heuristic, where popularity is the only cue considered by the agents to order the alternatives (for single attribute heuristics see Bagwell and Ramey, 1994; Hogarth and Karelaia, 2005). We call the order in which the alternatives are searched the *search path*. The order of search is the only point where social information influences agents that use the popularity heuristic to make decisions in cultural markets. The agents stop search when they encounter an alternative with utility higher than a threshold T_{pop}. The returns from examining one more alternative can be expressed as in (2). The sampling policy for an agent i is as follows:

$$\begin{aligned} &\text{if} \quad u_{ni} > T_{pop}, \quad \text{stop sampling} \\ &\text{if} \quad u_{ni} \leq T_{pop}, \quad \text{continue sampling} \end{aligned} \tag{4.4}$$

Although the popularity heuristic is behaviorally intuitive it is not closed-form optimal. A closed-form optimal solution would require that the agents infer the expected utility and

[6] This is where our model differs from informational cascade models (Banerjee, 1992; Bikhchandani et al., 1992) and agrees with more recent social influence models (Borghesi and Bouchaud, 2007; Hendricks et al., 2012).

the distributions of utilities for the alternatives available in the market on the basis of the current popularity information. In addition, a fully rational policy should be based on the assumption that other individuals are acting in a closed form optimal manner. Such an approach is computationally intractable for both artificial and human agents. The popularity heuristic is boundedly rational (Simon, 1955a). In both random search and the popularity heuristic, the threshold employed by the agents can be seen as a parameter in the model. In random search one can always tune the parameter to correspond to the optimal behavior. In the popularity heuristic, by contrast, no threshold level exists for which optimality is guaranteed. In the result section we will show that for the same threshold and cost of search the popularity heuristic outperforms random search. The fact that the heuristic outperforms an "optimally" derived strategy that can operate in the same environment already indicates that it is an effective strategy.

In both search rules, when all the alternatives in the market have been searched and no alternative has utility higher than T, the agent simply selects the alternative with the highest utility. Further, we assume that the agents can refrain from sampling any alternative when the cost of search is very high. In that case the agents are satisfied with the status quo outcome.

4.2.5 Assumptions

Here we make transparent and justify the modeling choices made in the previous section. First, we have assumed that agents employ a fixed threshold which in random search corresponds to the optimal stopping rule. This is a simplifying assumption, as in real life, people learn stopping rules gradually from experience in previous realizations of the market. Conlisk (2003) has shown how agents in random search problems can gradually converge to the optimal threshold by following a simple learning process.

Second, we have assumed that the agents can always recall alternatives that they have sampled in the past. This assumption is well suited for the markets of cultural products as the agents can in almost all cases find music albums, books or papers that they have listened to, read, or heard of in the past. Internet marketplaces and search engines have contributed a great deal to making previously examined cultural products available at any time.

Third, we have assumed that agents use only the popularity information to order the alternatives. In some scenarios this assumption may reflect actual consumer behavior (for example consider most read lists that are often widely used in online press), but in many cases agents have access to more information and they can use it to order the alternatives. Our model can be seen as a boundary case of an ordered search model in which several pieces of information are processed (e.g. Armstrong et al., 2009; Analytis et al., 2014). This strong assumption allows us to focus on the interplay between sequentiality in choice and diversity of preference.

Fourth, we assumed that the products do not bear a price tag and that can be consumed by an infinite number of agents. Indeed, in terms of production, cultural products are nonrivalrous information goods with a potentially unlimited supply. Clearly, after publishing a scientific article a great number of scientists can read it without increasing the marginal cost of production. If a universally best alternative existed, such as for example an outstanding online course in statistics, everybody would be better off by consuming that alternative without restricting the supply and increasing the price for other individuals. Most cultural products are excludable goods; That is, a price can be assigned to them. However, for many cultural products, for example movies, music albums and books, there tends to be a small price dispersion and in some cases there is even a predefined price (as in cinema tickets) for an entire product category. Yet sometimes, they are supplied freely by producer choice, as for example in the case of YouTube videos, massive online courses, or designer products distributed under a creative commons license. In the latter case cultural products can be categorized as pure public goods.

This brings us to the last major assumptions of the model. We have assumed that the agents have diverse preferences. The research on recommender systems demonstrates that this is clearly the case in practically all markets for cultural products. In fact, the degree of diversity in preferences varies in different cultural markets. Note that one could also assume variability in the stopping thresholds employed by the agents, but this would have only a marginal contribution to the results. Finally, we have opted to describe the subjective and objective utility components with a normal distribution for its neat aggregation properties and its plausibility.

4.3 Results

We simulated markets consisting of 1,000 agents that decided sequentially which of 100 alternatives to choose. For each market we drew the objective utilities u_o of the alternatives from a normal distribution with mean equal to 0 and variance equal to σ_o^2. For each individual agent in the market we drew the subjective utilities u_s of the alternatives from a different normal distribution with mean equal to 0 and variance equal to σ_s^2. For all the analyses reported in the results section we assumed that agents following random search employ the optimal stopping threshold. Then, we fixed the threshold parameter of the popularity heuristic to correspond to the optimal threshold for random search. We chose that implementation for two reasons. First, even if we were to test slightly different threshold rules, the main results of the simulation in terms of the emerging properties of the market would be very similar. It will soon become clear that the results are primarily driven by the ordering of the alternatives rather than by the stopping threshold. Second, in this way we can directly compare the welfare outcomes of the two models at the individual and aggregate level.

We systematically varied three parameters in our simulations. First, we varied the contribution of the objective utility u_o and the subjective utility u_s in the overall utility of the alternatives available in the market. This was done by setting the objective variance σ_o^2 equal to d, where $d = \{0, 0.1, 0.2, ..., 1\}$ is a diversity vector, and the variance of the subjective utility for each individual agent σ_s^2 equal to $1 - d$. Second, we varied the cost of search c that agents have to pay when sampling an alternative. We tested five possible costs, $c = \{1/2^3, 1/2^4, 1/2^5, 1/2^6, 1/2^7\}$, which correspond to the thresholds $T = \{0.78, 1.15, 1.47, 1.76, 2.03\}$. In this way we covered a large space of plausible costs encountered in actual markets for cultural products. When $c = 1/2^3$, the cost of searching one more alternative is a significant proportion of the average utility eventually enjoyed by the agents, whereas when $c = 1/2^7$, the cost of search is negligible. We opted to use the cost consistently on the x axis rather than the threshold although they can be used interchangeably. Note that in the first part of the Result section only the thresholds need to be known, whereas for the welfare analyses both the thresholds and the costs of search need to be known. Finally, we varied the decision-making process — random search or the popularity heuristic. Overall, this resulted in $11x5$ possible market environments for random search and the popularity heuristic, respectively. To account for randomness in the whole process, we ran each condition 100 times. In all the analyses conducted, except when we focused on unpredictability, we generated all the alternatives in the environment anew.

In the following sections we first examine how different search rules determine the emerging properties of the market, such as the market-share inequality, the outcome unpredictability and the relation between objective utility and popularity and then we turn to the welfare analysis. We wrote the code for running the simulation, analysing the results, and producing the figures in R language (Ihaka and Gentleman, 1996), and it is publicly available at the website: https://github.com/pantelispa/popularityHeuristic. The exact dynamics of the original simulation can be reproduced in https://hstojic.shinyapps.io/popularity/. We programmed the interactive environment with the Shiny application. Note that the application requires a strong computational system to run properly. In some systems it takes up to a minute per condition to load the original data of the simulation.

4.3.1 Utility, popularity and inequality

Let us first examine the aggregate behavior implied by each of the search rules. The popularity heuristic markets are characterized by evident rich-get-richer dynamics. The decisions of the first agents shape the search path that will be followed by agents deciding after them. Over time the search path tends to stabilize, which implies that the agents examine the alternatives in approximately the same order and tend to choose from the same subset. When the agents employ a lower threshold they examine a smaller subset of alternatives and swarm on fewer of them. In contrast, in random search markets for each agent each alternative that satisfies the threshold stands an equal chance of being selected ($1/k$, where k is the number of alternatives crossing the threshold). Thus, a lower threshold implies that more alternatives

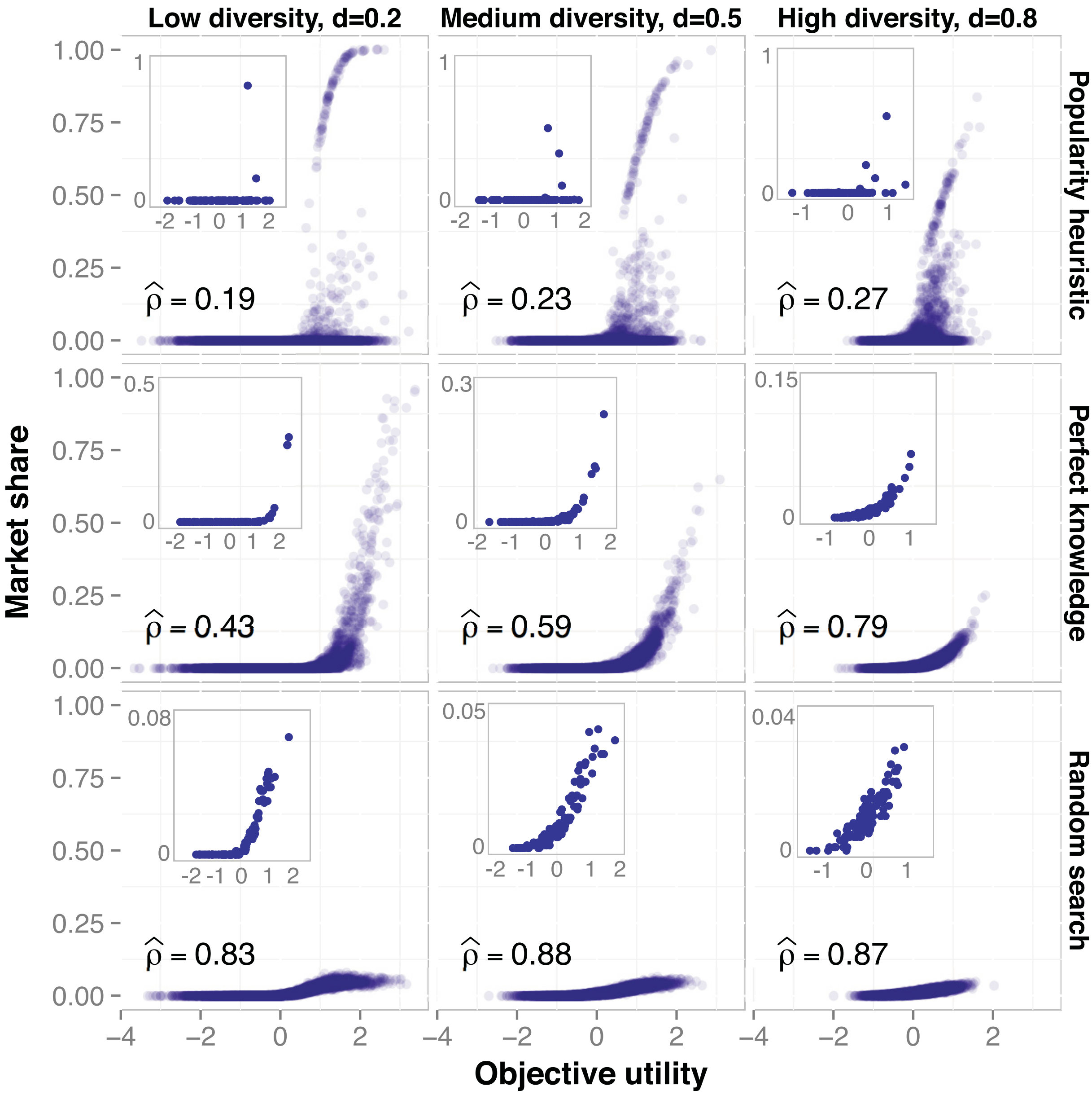


Figure 4.1: Up and botton row: The market share of all 100 alternatives in all 100 repetitions of the popularity heuristic and random search rule for the conditions with cost of search $c = 0.125$ and diversity of preferences $d = \{0.2, 0.5.0.8\}$. Middle row: The perfect knowledge condition where both processes converge for $c = 0$. The inset plots depict the results from a randomly drawn single run of the simulation. In the popularity heuristic condition, the objective utility of the alternatives is only weakly related to the obtained market shares at the end of the simulation. In contrast, in random search the market share is a very accurate predictor of the utility of the alternatives. We report the Pearson correlation in all the conditions.

stand an equal chance of being selected. Inversely, a higher threshold means sampling more broadly but choosing more selectively.

Note that for cost of search 0, both search processes converge to a perfect knowledge market (Figure 4.1 middle) since the agents sample all the alternatives regardless of the search rule. Then, a convex relationship between objective utility and market share exists that is especially pronounced when the subjective utility component is small. Such a market meets the assumption put forward by Rosen (1981). However, as the cost of search increases (stopping threshold decreases) the two processes diverge leading to distinct aggregate patterns.

At the end of the simulation under the popularity heuristic, many alternatives with high objective utility are never sampled or are sampled by very few individuals (see the inset graphs in Figure 4.1). Clearly, there is a great deal of luck involved at the first few steps of an emerging market, and having a high objective utility is necessary but not sufficient for succeeding in the market. Although it is possible to predict failure it is impossible to predict success. In contrast, the market share captured by the different alternatives increases gradually as a function of the objective utility component of the alternative in random search and quite steeply in the perfect knowledge condition. In the popularity heuristic market the objective utility is only weakly correlated with the obtained market share, while in random search more alternatives are choosen at least once and the correlation is much stronger.

Like in Salganik et al. (2006), we used the Gini coefficient to measure the inequality. The coefficient is defined as $G = \sum_{i=1}^{n} \sum_{j=1}^{n} \mid m_i - m_j \mid / 2n \sum_{k=1}^{n} m_k$. The term m_i stands for the market share of an alternative that is defined as $m_i = d_i / \sum_{l=1}^{n} d_l$, where d_i is the number of times the alternative was selected so far and d_l is the sum of choices made until decision maker l. The results for all the 110 conditions are shown in Figure 4.2. Lower costs of search (higher stopping thresholds) imply a lower Gini coefficient for the popularity heuristic, whereas they have the opposite effect for random search. Note that as the cost of search decreases (stopping threshold increases) the Gini coefficients of the two processes gradually converge to those of the perfect knowledge markets, which are $G = \{1, 0.98, 0.96, 0.93, 0.9, 0.85, 0.8, 0.72, 0.62, 0.46, 0.18\}$

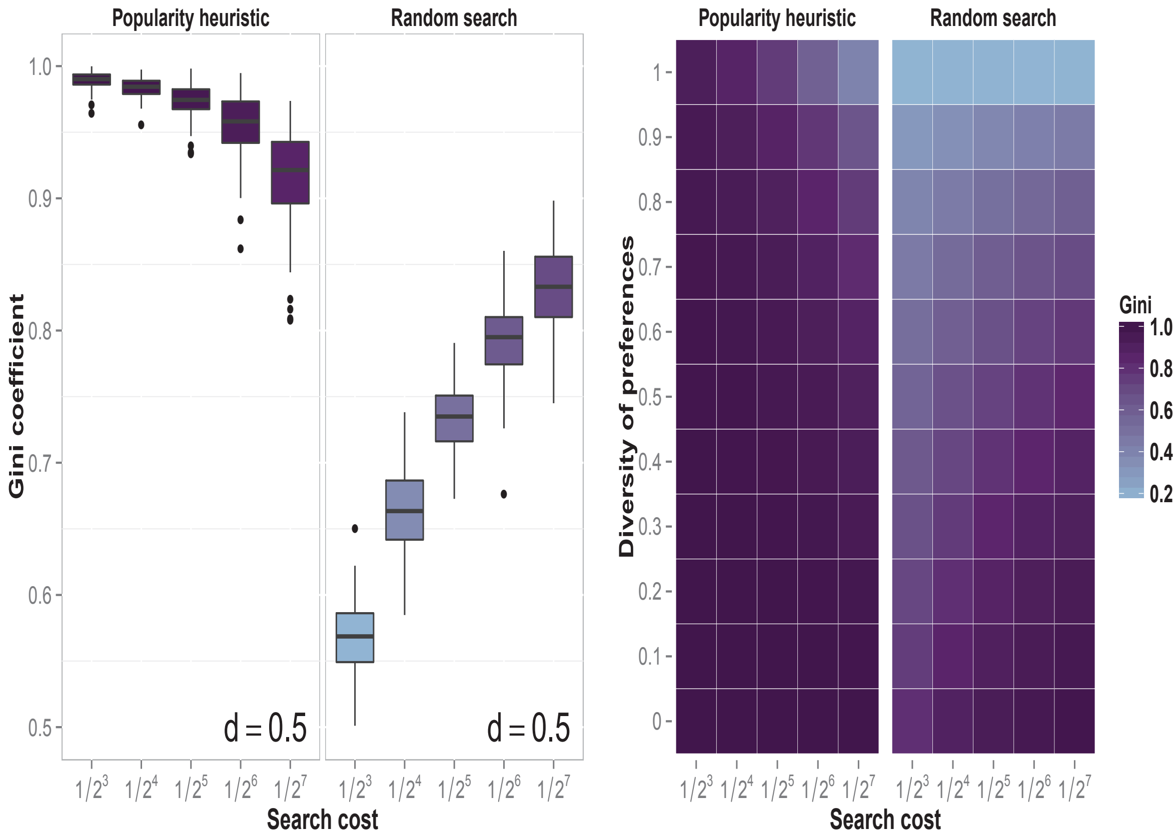


Figure 4.2: Left: the average Gini coefficient in a market with $d = 0.5$. In the popularity heuristic market a higher search cost (lower stopping threshold) leads to more inequality in the market. In contrast, in the random search market it implies a more egalitarian distribution. As the cost of search drops the Gini coefficient of the two processes converges to that of a perfect knowledge market. Right: the average Gini coefficient for all 55 popularity heuristic and random search markets.

for diversity of preferences $d = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, respectively.

4.3.2 Outcome unpredictability and luck

How accurately could we predict the popularity of the alternatives if we were able to rerun the entire market with the same initial conditions? To answer this question we generated 10 different distributions of objective utilities, which we refer to as worlds, and simulated each of these worlds 10 times every time with new agents. In this way we balanced the variability that could be caused due to the environment — consider a randomly generated world with an extraordinarily good alternative — ,with the need for additional replications of the same environment. Following Salganik et al. (2006) we define the unpredictability

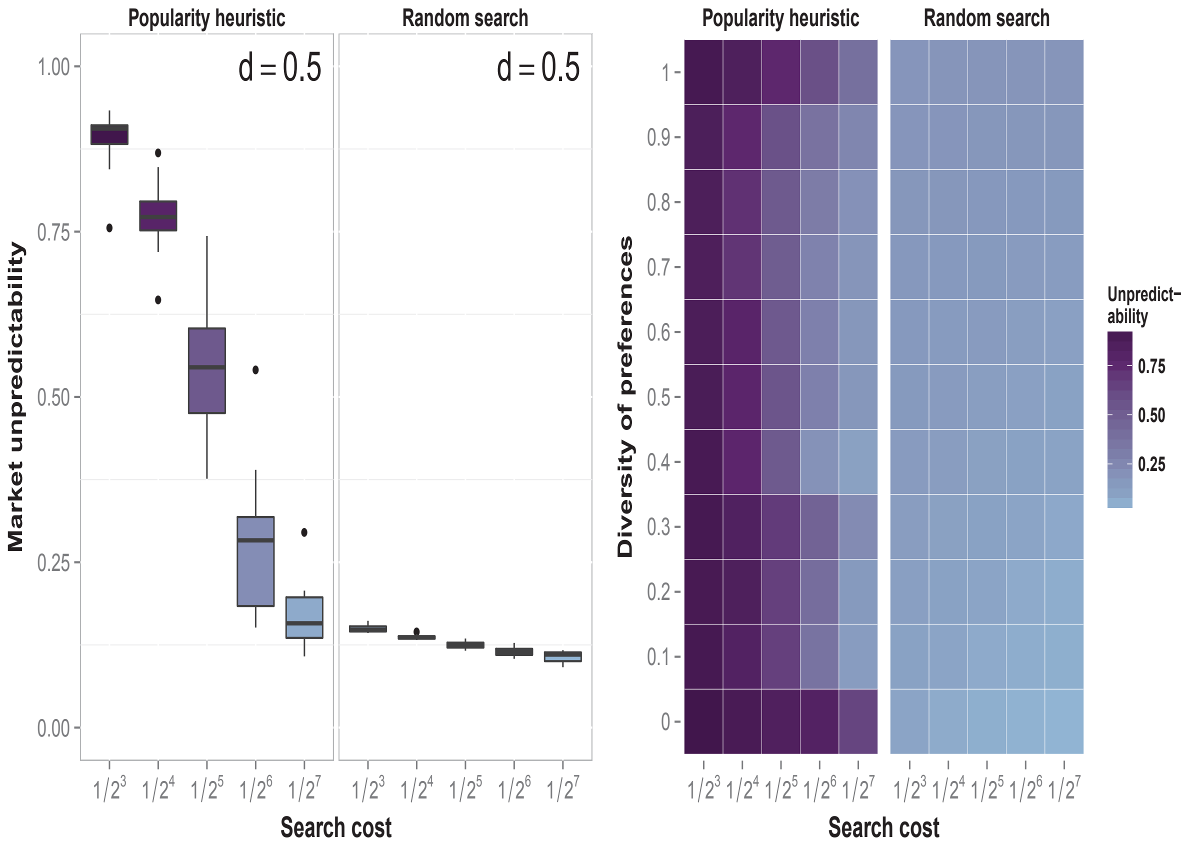


Figure 4.3: Left: The unpredictability in the market as a function of the cost of search when $d = 0.5$. For both the popularity heuristic and random search markets, the unpredictability in the market decreases as a function of cost of search, converging to 0, which corresponds to the perfect knowledge market. The convergence is much steeper in the popularity heuristic market were for high cost of search the market is very unpredictable. The variability of the unpredictability coefficient is large in the popularity heuristic market, whereas it is negligible in the random search market. Right: the unpredictability coefficient for all 55 conditions in the popularity heuristic and random search markets.

related to a specific alternative as $v_i = \sum_{j=1}^{W} \sum_{k=j+1}^{W} \mid m_{ij} - m_{ik} \mid / \binom{W}{2}$, where m_{ij} is the market share of an alternative i, in world j and W is the number of worlds. We then defined the overall unpredictability in an economy as $V = \sum_{i=1}^{n} v_i/2$. Following this formula the results have an intuitive interpretation that holds for any number of alternatives. Unpredictability 1 corresponds to the case where the entire market share goes to a different alternative in every single replication (world) of the simulation. Unpredictability 0 corresponds to the case where the market share is distributed identically in all the replications (worlds) of the simulation.[7]

[7]Salganik et al. (2006) and Salganik and Watts (2008, 2009) defined the overall unpredictability as $V = \sum_{i=1}^{n} v_i/S$ where S is the number of alternatives in the market (in their case, songs). Krumme et al. (2012) and Borghesi and Bouchaud (2007), who have developed models that can capture the dynamics of the Music

As depicted in Figure 4.3 both search processes converge towards unpredictability 0 as the cost of search decreases (threshold increases), which corresponds to a perfect knowledge market. The popularity heuristic markets are very unpredictable for high costs but become fairly predictable for low costs. In contrast, the unpredictability reduces at a slow pace in the random search markets. Similar results are obtained for all values of d. In the popularity heuristic markets unpredictability as a function of diversity of preferences is bimodal — the lowest levels of unpredictability are obtained for intermediate levels of d, while it increases as d goes to 0 and 1. Reasons for higher unpredictability in extremes differ. As d goes to 0, preferences become correlated and first movers are dictators of the market, but what agent is the first mover is random from world to world. On the other hand, as d goes to 1, preferences differ a lot and each world is driven by idiosyncratic agents preferences which are world specific. In contrast for intermediate levels of uncertainty agents search a bit longer and for varying lengths down the same search path since their preferences differ to some extent. This leads to the accumulation of more valuable social information. Different agents eventually often choose the same objectively better alternatives. Another point worth mentioning is that while both environments are uncertain, random search markets are predictable in the long run whereas the popularity heuristic markets are not. The popularity heuristic markets are non-ergodic, which means that small events have irreversible consequences in the course of history, whereas the random search markets are ergodic random processes that converge to approximately the same results for large samples (for more details on non-ergodicity see Arthur, 1989).

4.3.3 Welfare analysis

In Figure 4.4, we present the average utility obtained in 55 different popularity heuristic conditions. First, note that in random search the average utility in the market does not depend on the diversity of preferences in the population but merely on the search cost variable.

Lab experiment also used this formula in their studies. However, according to this formula the maximum degree of unpredictability depends on the number of alternatives available in the market leading to results that cannot be compared across studies with a different number of alternatives.

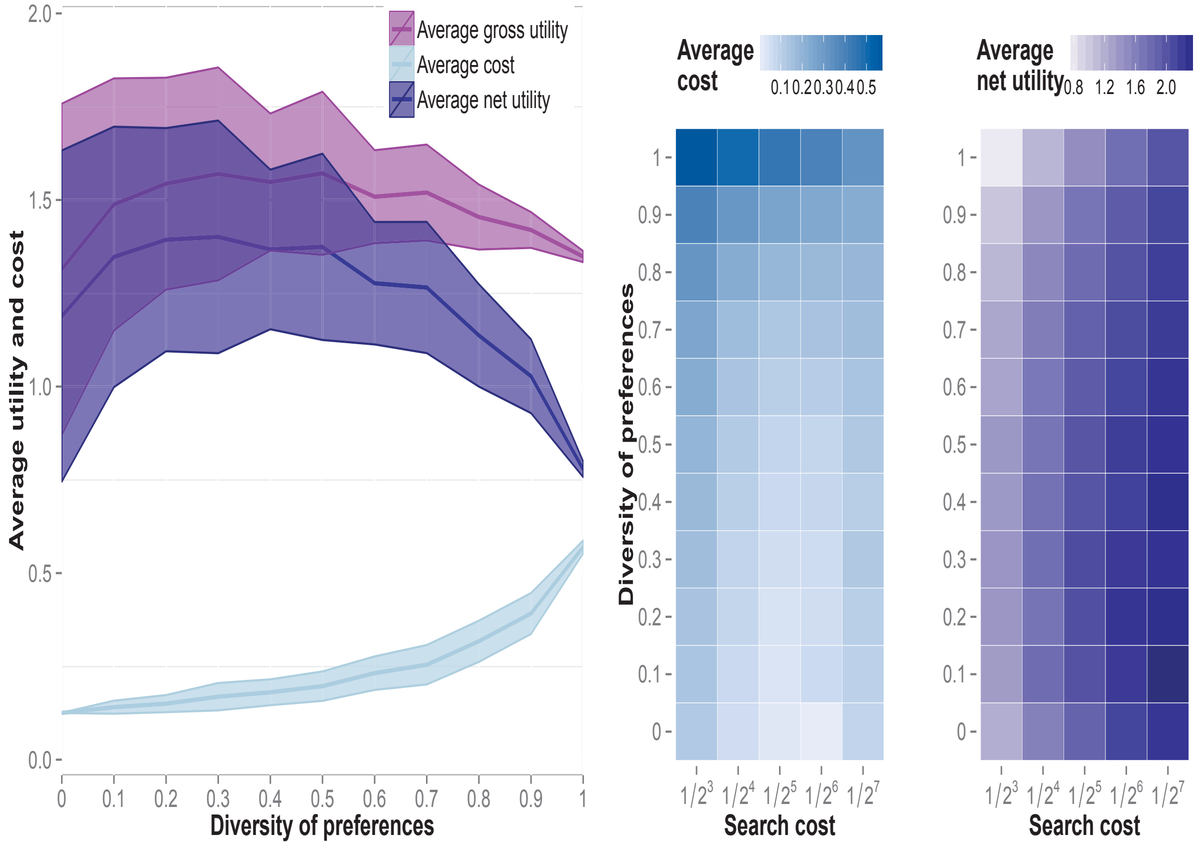


Figure 4.4: Left: The average gross and net utility of the selected alternative and the average total cost of search for search cost $1/2^3$. The ribbons represent the variability in 100 repetitions of the simulation. The average net utility in the market is the highest for intermediate levels of diversity of preferences. Right: the average net utility and the average total cost of search in the 55 popularity heuristic conditions. The graph on the left side corresponds to the leftmost columns of the heatmaps on the right side.

As a result, there are only five relevant conditions that are perfectly equivalent, in terms of average utility and cost of search, with the popularity heuristic conditions with $d = 1$. When $d = 1$ the environment is recreated randomly for each agent. Thus, even though the agents search according to the popularity order, they essentially behave the same as agents that search randomly. In these conditions the agents incur the highest average total cost of search and earn the lowest average net utility from the market. To find out the extent to which the popularity heuristic outperforms random search in various diversity of preferences conditions, the reader can simply compare them with the $d = 1$ condition. Also note that in a perfect knowledge market (c = 0), the agents always consume the alternative with the highest utility in the market, which on average has utility $u = 2.51$.

We find that as d decreases, the cost of search steadily declines, as well. When all the agents have exactly the same preferences, the first agent pays total search cost kc (where k stands for the number of alternatives sampled until the agent encounters an alternative with utility higher than T). All the remaining agents simply imitate the decision made by the first agent and pay c only once. Following the terminology introduced by Page (2006), the process is initial outcome dependent. All the agents deterministically follow the choice of the first individual. The surprising result is that although the average agent in markets with $d > 0$ incurs a higher average cost of search than the average agent in a market with $d = 0$, this loss is outweighed by the benefits of choosing alternatives with higher gross utility. For high search costs, for example, intermediate levels of d lead to the highest average net utility. As the cost of search declines, however, the amount of beneficial diversity of preferences decreases. Consequently, the differences in the average utility among the conditions decline, as can be seen in the rightmost columns of the net utility heatmap (Figure 4.4). Finally, because agents tend to follow first agents as d decreases and their preferences become more correlated, and there is luck in who moves first, a lower d implies more variability in the average utility in the market. This is illustrated by the ribbons in the left panel of Figure 4.4.

This analysis revealed an unexpected result. Although the costs of search are lowest in markets with no diversity of preferences, the average net utility is higher in economies with moderate diversity. What is the process that leads to an increase in the average obtained utility in the market in such conditions? To gain insight into the process, we examined the evolution of net utility in the market. We plotted the average net utility enjoyed by the agents $\sum_{i=1}^{n} u_i/b$ in sequential blocks of 25 agents for search cost $1/2^3$ (left panel in Figure 4.5). This analysis reveals that in contrast to a market without diversity where the average utility remains constant, in markets with some diversity the average utility continues to increase in the first agent blocks. As discussed above, in a market with no diversity whatsoever, everybody is satisfied with a decision made by the first individual. Occasionally this turns out to be an excellent alternative, but more often it is located just off the satisficing threshold. In

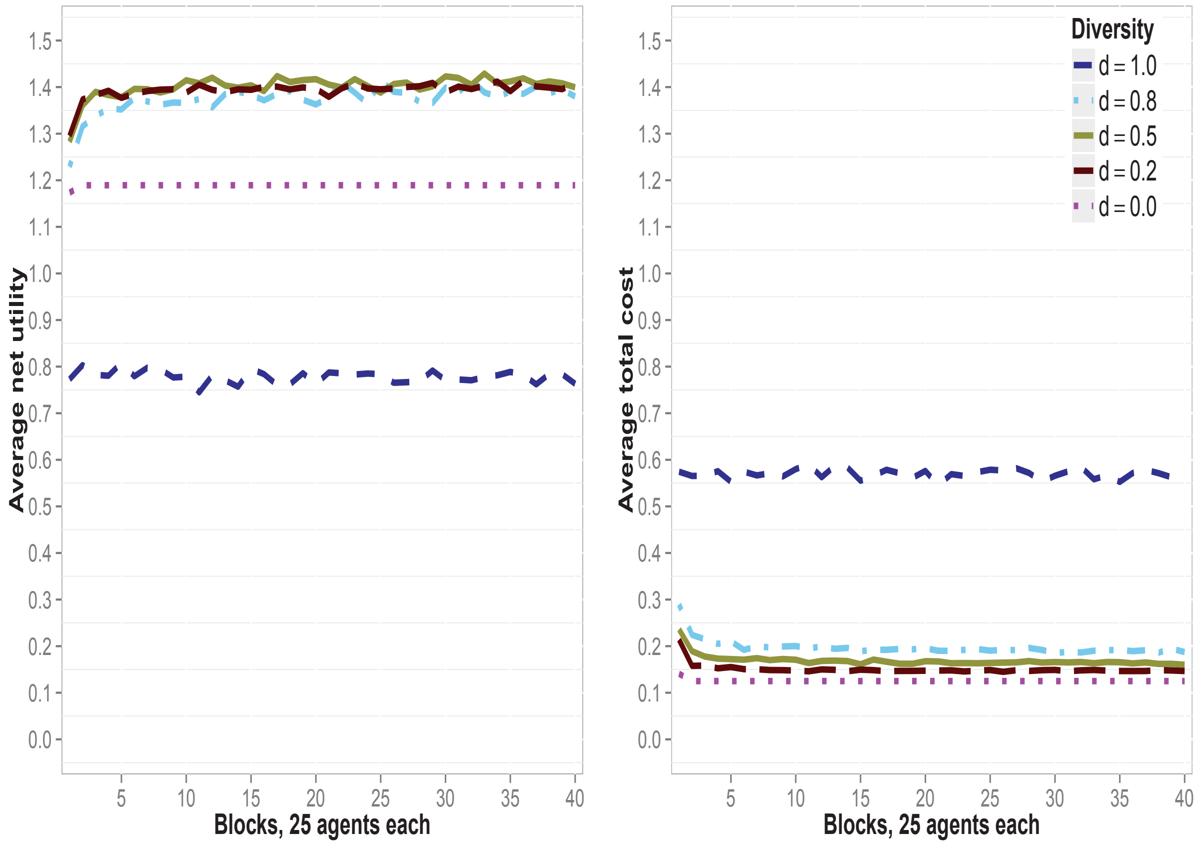


Figure 4.5: The average net utility in the market $\sum_{i=1}^{l} u_i/25$ (left panel) and the average total cost of search $\sum_{i=1}^{l} k_i c/25$ (right panel) where l was set to blocks of 25 agents, for search cost $1/2^3$. When $d < 1$, the average net utility in the market tends to increase as more agents make their choices, but with diminishing returns with each additional agent. The average total cost in the market has the opposite pattern: it decreases with diminishing returns with each additional agent. The largest increase (decline, respectively) occurs in the very first agent blocks.

contrast, in markets with some diversity of preferences, alternatives with a higher objective utility component tend to be gradually pulled toward the beginning of the search order. Because preferences are not perfectly correlated agents will search a bit further until they encounter a satisficing alternative. The search path followed by the agents gradually improves in terms of the encountered objective utilities after the decisions of several first agents. Consequently, alternatives with higher objective utility components, which in some cases are even higher than the threshold, tend to conquer the first spots in the search path. Thus, the agents deciding later in time have a more useful popularity signal, require less search to find an alternative that satisfies Condition (4), and tend to settle on much better alternatives.

In essence, some diversity of preferences allows the agents to benefit from the informational externalities, that are part of the sequential social interaction. By searching and deciding, the agents unintentionally provide useful information to the individuals deciding after them. The increase of the average net utility in the economy is observed for all diversity of preferences conditions, but it is strongest when the cost of search is high (or the satisficing threshold is low). The reason behind this is clear: With high search cost the agents search relatively shallowly, leaving a greater potential for further improvement by reordering the alternatives in the search path.

4.4 Discussion

Sixto Diaz Rodriguez, a talented folk/rock musician from Detroit for who many thought destined for great success, released two promising albums at the beginning of the 1970s. To the disappointment of his producer, his albums sold only a few copies in the United States and most other countries in the world. However, Rodriguez's music was an absolute hit in South Africa, Botswana, Australia and New Zealand. Rodriguez himself was completely unaware of his superstar status in these countries until he was invited to tour in Australia in the late 1970s and in South Africa only in 1990s.[8] In the absence of counterfactual evidence, like the different regional markets in the story of Rodriguez, we could quickly conclude that differences in quality are enough to explain the success or failure in a market. The Music lab experiment and recent models of social influence have demonstrated that we have to take into account the social influences that lead to rich-get-richer dynamics in a market. With these effects exactly the same initial conditions can lead to different end-market outcomes. Agent-based simulations have the advantage that counterfactual worlds can be created very cheaply. This way, simulations can generate predictions for environments with conditions different from the data at hand.

Agents in our simulations followed a sequential search process (Simon, 1955a; DeGroot, 1970), a model with extensive applications in economics and decent experimental support (Rapoport and Tversky, 1970; Hey, 1987). We chose this model because it maps important characteristics of cultural markets better than the models used in all previous studies. This includes sequentiality both within agents and between agents, costly sampling and focusing on only a subset of alternatives due to the huge volume of the market. We introduced social information into the model through a popularity heuristic, whereby agents examined alternatives according to aggregate popularity information. Aggregate popularity information is a signal used widely in real life cultural markets — number of citations and book bestseller-lists being prime examples. Importantly, recent findings that popularity information influences the

[8]The story of Rodriguez was portrayed in documentary film "Searching for Sugar Man" which eventually earned the Oscar as the best documentary film.

sampling process rather than the decision process are in line with the usage of popularity information in our model (Krumme et al., 2012; Denrell and Le Mens, 2013). We used agent-based simulations to scale up from individual agents to the market outcomes. Our results might not have the generality of analytical solutions, but we did not have to simplify the model or commit to strict assumptions about our agents knowledge or information-processing capabilities, which allowed us to maintain a good match to cultural markets.

4.4.1 Luck and dynamics of success in cultural markets

Stochastic process models suggest that the success of an alternative in the market is a matter of pure luck or merely early arrival in the market (Price, 1976). Differences in quality are completely inconsequential with respect to success. In contrast, in models in which agents have full knowledge about utilities, the eventual market shares of the alternatives are completely determined by utilities (Rosen, 1981). Salganik et al.'s (2006) experiments indicated that both quality and luck play a role. Our model pins down an exact cumulative advantage mechanism by which quality (objective utility) and luck play out. Success in the market depends on both the utility of an alternative and the decisions of a few influential decision makers in the market. In our model, conquering one of the first positions in the search order confers a significant advantage. Alternatives with a high objective utility component stand a better chance of reaching these positions, which through social interactions may attract many more consumers over time. Hendricks et al. (2012) reached the same result with their model, but we show that it holds in much richer setting with within agent sequentiality. This setting led us to some novel insights on the inner workings of the market. For example, in our model the decision makers unavoidably remain unaware of the existence of many high-quality alternatives — the alternatives that will be pulled to the first position of the popularity list are only a subset of all the high quality alternatives and it is probable that the alternative with the highest objective utility is not among them. This provides a rationale for firms competing to place their products along the consumers' search paths (e.g. Armstrong et al., 2009; Bagwell and Ramey, 1994). A more important benefit of using this setting was that it allowed us to analyze welfare outcomes for the market, which we discuss in details in the following section.

4.4.2 Exploration, imitation and diversity

Search models exemplify the trade-off between exploiting the best alternative found thus far and further exploring with the hope of identifying better ones (March, 1991). Exploration is based on trial and error and entails a high cost for individual decision makers. Thus, in sequential decision-making environments self-interested decision makers are often better off by imitating the choices of individuals who have already incurred the costs of exploration. As an example, consider our model in the conditions with no diversity of preferences. Although imitation makes sense at the individual level, it deprives the group from additional information that could have been collected by individual explorers (Rogers, 1988). This is why in eusocial species such as bees and ants, in which individuals have evolved to maximize the fitness of the entire group, the members of the group occasionally explore new alternatives for hive locations or better paths to reach food sources.[9] Information cascade and herding models, developed by economists portray the shortcomings of imitation and social influence and stress the potential losses of collective welfare (e.g. Bikhchandani et al., 1992). In contrast, social learning studies tend to focus on the effectiveness of imitation strategies at the individual level (e.g. Laland, 2004; Richerson and Boyd, 2008; Rendell et al., 2010). Depending on the quality of feedback, and the pay-off structure in the environment, imitation can lead to either passable or very good results at the aggregate level.

Even in environments where agents are self-interested there are mechanisms that can keep imitation forces in check and reap some of the benefits of the independent collection of information. For example, under different sets of conditions a natural equilibrium evolves between explorers (or producers) and the imitators (or scroungers) (Conlisk, 1980; Rogers, 1988; Kameda and Nakanishi, 2002). Further, in some search environments, barriers to

[9] Their algorithms are so effective that they have been emulated in operations research to derive solutions to combinatorial optimization problems (Dorigo and Gambardella, 1997; Karaboga and Basturk, 2007)

communication in the form of a sparser communication network among agents turn out to be beneficial at the collective level. They encourage individuals to explore more, supplying in this way useful information to the group (Lazer and Friedman, 2007; Mason et al., 2008; Fang et al., 2010). We believe that the interplay between diversity and costly sequential search provides another such example. Because preferences differ and agents can find out the quality of alternatives, they sample some alternatives rather than unconditionally imitating the choice of agents deciding earlier. The popularity information influences the order of search, but crucially agents still sample some alternatives, which ensures that some new information is stored in their choices. Over time this behavior improves the search path and increases the welfare of the collective (see Figure 4.5). Note, however, that too much diversity makes the signals less informative since preferences become completely uncorrelated, which then leads to a decrease in the aggregate welfare.

Similar counter-intuitive beneficial effects of diverse preferences have been shown in an earlier sequential model of social learning (Goeree et al., 2006) and a recent publication on the dynamics of the academic reviewing process (Park et al., 2014). In line with our account, in these models diversity incentivizes agents to use individual information providing in this way useful information to the group. In all these three examples diversity can be seen as a beneficial hurdle to communication that helps to internalize the strong informational externalities. This mechanism is distinct from alternative accounts for the bebeficial role of diversity in groups that are based on complementarities in the problem solving capacities of the agents in a group (Clearwater et al., 1991; Hong and Page, 2004), or complementarities in the information possessed by diverse members of a group (Conradt et al., 2013; Davis-Stober et al., 2014).

Overall, in our study socially informed search led to an improvement in the overall welfare as compared to random search. As illustrated in Figure 4.1, random search generates much more useful information, but it does not allow an individual agent to save any costs. Thus, an agent would prefer to be in a market of random searchers, but would choose using the popularity heuristic. In our model lower costs of search drastically changed the performance of the popularity heuristic, led to much better aggregate welfare outcomes and lessened the extent to which diversity can be beneficial for the collective.

4.4.3 Revisiting the Music lab experiment

The Music Lab illustrated how history can be rerun in the lab (Salganik et al., 2006). The authors conducted the experiment with clearly stated hypotheses, but without specifying a precise individual decision-making model beforehand. They suggested that agent-based models could complement their research and generate predictions for other possible conditions of the experiment — a different number of alternatives, or an evolving market (Salganik and Watts, 2009). As already discussed in the Introduction, two computational models have attempted to capture the decision-making processes of the participants in the experiment (Borghesi and Bouchaud, 2007; Krumme et al., 2012). Fitting the parameters of the models to the data they can reproduce the findings of the experiment.[10] Although inspired by Salganik et al.'s (2006) experiments, our model was not designed to reproduce its exact dynamics. Instead it was conceived to generate novel insights about the role of social influence stemming from a well-founded theoretical framework and to highlight the welfare side of the Music Lab market which was not studied in the experiments or in the follow-up modeling attempts. Yet, our model generates clear-cut hypotheses about the direction of change of all the main measures of the Music Lab experiments. Thus, it could serve as a guide for future experimentation or the collection of field data. For example, our model would predict that when the number of alternatives increases the inequality and unpredictability in the market should also increase. Similarly, it would predict more inequality and unpredictability in markets in which the agents have more homogeneous preferences. In the future we believe that a more constrained experimental design, in which the number of alternatives to be selected is pre-specified, would further help to uncover the individual decision-making processes.

[10]Note that we could fix several parameters of the model, such as the number of alternatives selected by each agent, the diversity and the satisficing threshold in order to reproduce as closely as possible the aggregate dynamics.

4.4.4 Managing social influence

Creators, producers and other individual actors in the markets for cultural products are primarily interested in maximizing their own benefits from the market. In contrast, the designers of online cultural marketplaces (such as Amazon), or public organizations might be primarily interested in maximizing the welfare of the entire agent population. To this end, the designers may attempt to manage social influence by channeling information selectively to the agents in the decision-making sequence. An early example of large-scale social influence management is provided by the work of Wu and Huberman (2008) who examined the social network Twitter and studied the impact of different post ordering algorithms on the total number of clicks. How should the social influence be structured if the goal is to maximize the sum of the utilities of the participants in the market? Our model reveals that agents deciding early on in the sequence have a large impact on the average utility in the economy. In addition, agents who act independently of each other supply more information to the agents deciding after them. Hence, a possible approach would be to give to n agents in the beginning of the sequence the task of randomly exploring the alternatives. Then, the remaining agents would follow the search path shaped by those first agents. Such a tentative analysis could be easily conducted within the current framework or its extensions in the future.

4.4.5 Improving the external validity

Hitherto, we have opted for a minimal and psychologically plausible process model that is still able to capture the stylized facts of the markets for cultural products. Clearly, in many actual markets the agents decision-making processes are psychologically richer than our account. For example, how would social influence play out if agents had access to further information such as the average rating from other users in addition to the popularity information?[11] Similarly, how would awareness or recognition of some alternatives (e.g. Narayana and Markin, 1975; Goldstein and Gigerenzer, 2002; Fleder and Hosanagar, 2009)

[11] Characteristically, Netflix implements a ranking algorithm that combines popularity with other non-social information to order the alternatives for its users.

change an individual's search and choice strategy and as a result the collective behavior? Or, how would the collective patterns change if agents accessed information locally from their group of friends and acquaintances (Pachur et al., 2005) rather than globally from a general bestseller list? Our model provides a framework within which such psychological insights and their implications for aggregate welfare could be studied in the future. Extensions of the model could take into account other sources of information, friendship networks or product awareness. Such extensions would also be crucial for further understanding and then managing social influence in real-world marketplaces.

4.4.6 Concluding remarks

To understand most real-world phenomena, like the initial failure and the eventual triumph of Sixto Diaz Rodriguez, it does not suffice to have an accurate model of individual behavior that disregards social interaction. We also need to understand how social influence alters individual behavior and leads to collective macroscopic patterns. The study of collective behavior has been gaining momentum over the last few decades. New methodological possibilities such as the accessibility of big data (e.g. Wu and Huberman, 2007; Bentley et al., 2014), the possibility to conduct large-scale experiments over the Internet (e.g. Salganik et al., 2006), the rise of agent-based simulations (e.g. Schelling, 1978), or the coupling of experimentation with simulation methods (e.g. Moussaïd et al., 2013) have contributed to its development (also see Goldstone and Janssen, 2005; Goldstone and Gureckis, 2009).

In this paper we developed a parsimonious process model of collective behavior based on a psychologically plausible process model of individual decision-making that can capture the stylized facts observed in the markets for cultural products. We manipulated diversity of preferences and search costs and derived predictions about the expected inequality and unpredictability in the market, as well as the welfare of the participants in them, for the entire parametric space. The model can inform the decisions of many of the stakeholders participating in the markets: producers, for instance, betting on which music albums will succeed and figuring out how to distribute the risk among them or marketeers planning to launch a campaign for a product that has not been selling well. It could be a tool in the hands of policy makers who need to design and justify scholarship schemes to support promising yet unknown artists. Moreover, it gives insights on how the markets could be designed to improve the average welfare for the consumers acting in them. Finally, the model can be used to generate new hypotheses for future large-scale studies such as the Music Lab experiment and it provides a framework within which additional psychological processes that determine the market dynamics might be investigated in the future.

4.5 References

Adler, M. (1985). Stardom and talent. *The American Economic Review*, 75(1):208–212.

Analytis, P. P., Kothiyal, A., and Katsikopoulos, K. (2014). Multi-attribute utility models as cognitive search engines. *Judgment and Decision Making*, 9(5):403–419.

Armstrong, M., Vickers, J., and Zhou, J. (2009). Prominence and consumer search. *The RAND Journal of Economics*, 40(2):209–233.

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal*, 99(394):116–131.

Bagwell, K. and Ramey, G. (1994). Coordination economies, advertising, and search behavior in retail markets. *The American Economic Review*, 84(3):498–517.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Bentley, R. A., O'Brien, M. J., and Brock, W. A. (2014). Mapping collective behavior in the big-data era. *Behavioral and Brain Sciences*, 37(01):63–76.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):992–1026.

Borghesi, C. and Bouchaud, J.-P. (2007). Of songs and men: a model for multiple choice with herding. *Quality & Quantity*, 41(4):557–568.

Brock, W. A. and Durlauf, S. N. (2001). Discrete choice with social interactions. *The Review of Economic Studies*, 68(2):235–260.

Chow, Y. S., Robbins, H., and Siegmund, D. (1971). *Great expectations: The theory of optimal stopping*. Boston, MA: Houghton Mifflin Boston.

Chung, K. H. and Cox, R. A. (1994). A stochastic model of superstardom: An application of the yule distribution. *The Review of Economics and Statistics*, 76(4):771–775.

Clearwater, S. H., Huberman, B. A., and Hogg, T. (1991). Cooperative solution of constraint satisfaction problems. *Science*, 254:1181–1183.

Conlisk, J. (1980). Costly optimizers versus cheap imitators. *Journal of Economic Behavior & Organization*, 1(3):275–293.

Conlisk, J. (2003). A note on convergence of adaptive satisficing to optimal stopping. *Journal of political economy*, 111(6):1353–1360.

Conradt, L., List, C., and Roper, T. J. (2013). Swarm intelligence: When uncertainty meets conflict. *The American Naturalist*, 182(5):592–610.

Davis-Stober, C. P., Budescu, D. V., Dana, J., and Broomell, S. B. (2014). When is a crowd wise? *Decision*, 1(2):97–101.

De Vany, A. (2004). *Hollywood economics: How extreme uncertainty shapes the film industry.* London, England:Routledge.

DeGroot, M. (1970). *Optimal statistical decisions.* New York, NU: McGraw-Hill.

Denrell, J. and Le Mens, G. (2007). Interdependent sampling and social influence. *Psychological Review*, 114(2):398–422.

Denrell, J. and Le Mens, G. (2013). Information sampling, conformity and collective mistaken beliefs. In *in M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), Proceedings of the 35th Annual Conference of the Cognitive Science Society.*, pages 2177–2182. Austin, TX: Cognitive Science Society.

Dorigo, M. and Gambardella, L. M. (1997). Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, 1(1):53–66.

Elster, J. (1989). *Nuts and bolts for the social sciences*. Cambridge England: Cambridge Univ Press.

Fang, C., Lee, J., and Schilling, M. A. (2010). Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization Science*, 21(3):625–642.

Fiedler, K. and Juslin, P. (2006). *Information sampling and adaptive cognition*. Cambridge England: Cambridge University Press.

Fleder, D. and Hosanagar, K. (2009). Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management Science*, 55(5):697–712.

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

Goeree, J. K., Palfrey, T. R., and Rogers, B. W. (2006). Social learning with private and common values. *Economic Theory*, 28(2):245–264.

Goldstein, D. G. and Gigerenzer, G. (2002). Models of ecological rationality: the recognition heuristic. *Psychological Review*, 109(1):75–90.

Goldstone, R. L. and Gureckis, T. M. (2009). Collective behavior. *Topics in Cognitive Science*, 1(3):412–438.

Goldstone, R. L. and Janssen, M. A. (2005). Computational models of collective behavior. *Trends in Cognitive Sciences*, 9(9):424–430.

Granovetter, M. and Soong, R. (1986). Threshold models of interpersonal effects in consumer demand. *Journal of Economic Behavior & Organization*, 7(1):83–99.

Hendricks, K. and Sorensen, A. (2009). Information and the skewness of music sales. *Journal of Political Economy*, 117(2):324–369.

Hendricks, K., Sorensen, A., and Wiseman, T. (2012). Observational learning and demand for search goods. *American Economic Journal: Microeconomics*, 4(1):1–31.

Hertwig, R., Herzog, S. M., Schooler, L. J., and Reimer, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(5):1191–1206.

Hey, J. D. (1987). Still searching. *Journal of Economic Behavior & Organization*, 8(1):137–144.

Hogarth, R. M. and Karelaia, N. (2005). Ignoring information in binary choice with continuous variables: When is less more? *Journal of Mathematical Psychology*, 49(2):115–124.

Hong, L. and Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46):16385–16389.

Ihaka, R. and Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3):299–314.

Kameda, T. and Nakanishi, D. (2002). Cost–benefit analysis of social/cultural learning in a nonstationary uncertain environment: An evolutionary simulation and an experiment with human subjects. *Evolution and Human Behavior*, 23(5):373–393.

Karaboga, D. and Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: artificial bee colony (abc) algorithm. *Journal of Global Optimization*, 39(3):459–471.

Krumme, C., Cebrian, M., Pickard, G., and Pentland, S. (2012). Quantifying social influence in an online cultural market. *PLOS ONE*, 7(5):e33785.

Laland, K. N. (2004). Social learning strategies. *Animal Learning & Behavior*, 32(1):4–14.

Lazer, D. and Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative Science Quarterly*, 52(4):667–694.

Lee, M. D. (2006). A hierarchical bayesian model of human decision-making on an optimal stopping problem. *Cognitive Science*, 30(3):1–26.

Leibenstein, H. (1950). Bandwagon, snob, and veblen effects in the theory of consumers' demand. *The Quarterly Journal of Economics*, pages 183–207.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1):71–87.

Mason, W. A., Jones, A., and Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, 137(3):422–433.

McCall, J. J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–126.

Merton, R. K. (1968). The matthew effect in science. *Science*, 159(3810):56–63.

Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, 1(2):226–251.

Moussaïd, M., Kämmer, J. E., Analytis, P. P., and Neth, H. (2013). Social influence and the collective dynamics of opinion formation. *PLOS ONE*, 8(11):e78433.

Muth, J. F. (1986). Search theory and the manufacturing progress function. *Management Science*, 32(8):948–962.

Narayana, C. L. and Markin, R. J. (1975). Consumer behavior and product performance: An alternative conceptualization. *The Journal of Marketing*, pages 1–6.

Newman, M. E. (2005). Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46(5):323–351.

Pachur, T., Rieskamp, J., and Hertwig, R. (2005). The social circle heuristic: Fast and frugal decisions based on small samples. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum*, pages 1077–1082. Citeseer.

Page, S. E. (2006). Path dependence. *Quarterly Journal of Political Science*, 1:87–115.

Park, I., Peacey, M., and Munafò, M. (2014). Modelling the effects of subjective and objective decision making in scientific peer review. *Nature*, 506(7486):93–96.

Price, D. d. S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27(5):292–306.

Rapoport, A. and Tversky, A. (1966). Cost and accessibility of offers as determinants of optional stopping. *Psychonomic Science*, 4(1):145–146.

Rapoport, A. and Tversky, A. (1970). Choice behavior in an optional stopping task. *Organizational Behavior and Human Performance*, 5(2):105–120.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., Fogarty, L., Ghirlanda, S., Lillicrap, T., and Laland, K. N. (2010). Why copy others? insights from the social learning strategies tournament. *Science*, 328(5975):208–213.

Reutskaja, E., Nagel, R., Camerer, C. F., and Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *The American Economic Review*, pages 900–926.

Richerson, P. J. and Boyd, R. (2008). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.

Rogers, A. R. (1988). Does biology constrain culture? *American Anthropologist*, 90(4):819–831.

Rosen, S. (1981). The economics of superstars. *The American economic review*, 71(5):845–858.

Salganik, M. J., Dodds, P. S., and Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762):854–856.

Salganik, M. J. and Watts, D. J. (2008). Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social Psychology Quarterly*, 71(4):338–355.

Salganik, M. J. and Watts, D. J. (2009). Web-based experiments for the study of collective social dynamics in cultural markets. *Topics in Cognitive Science*, 1(3):439–468.

Schelling, T. C. (1978). Micromotives and macrobehavior.

Simon, H. A. (1955a). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.

Simon, H. A. (1955b). On a class of skew distribution functions. *Biometrika*, pages 425–440.

Telser, L. G. (1973). Searching for the lowest price. *American Economic Review*, 63(2):40–49.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of small-worldnetworks. *Nature*, 393(6684):440–442.

Wu, F. and Huberman, B. A. (2007). Novelty and collective attention. *Proceedings of the National Academy of Sciences*, 104(45):17599–17601.

Wu, F. and Huberman, B. A. (2008). Popularity, novelty and attention. In *Proceedings of the 9th ACM conference on Electronic commerce*, pages 240–245. ACM.

Yule, G. U. (1925). A mathematical theory of evolution, based on the conclusions of dr. jc willis, f. r. s. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 213(402-410):21–87.

Chapter 5

General discussion

Making sense of patterns of collective behavior and predicting aggregate behavioral measures is essential for acting effectively in economic and social contexts. By definition, psychology is the science of behavior, yet until today there have been relatively few occasions where psychological theories have been employed to understand aggregate behavioral phenomena. However, things are changing. Recent scientific developments such as the rise of behavioral economics and computational social science have transformed the way in which we understand the ties between individual and aggregate behavior. Scientists trained in psychology can now better than ever leverage the comparative advantages of psychological process models to tackle problems of aggregate behavior encountered in the social sciences.

In the three projects presented in this dissertation, I have advanced novel process models and examined their implications for individual and aggregate behavior. In chapter two, I introduced a psychological search and stopping theory of choice that integrates two milestone models of rational choice: one-shot multi-attribute utility theory and optimal stopping theory. In chapter three, I presented the results from two experiments designed to study how people revise their opinions when they are exposed to the opinion and confidence of another individual. Then, on the basis of these experimental observations, I delineated a decision tree that describes individual behavior under social influence. The model was scaled-up from the individual to the aggregate level by way of an agent-based simulation. Finally, in chapter four, I presented a psychologically plausible model of choice under social influence which, at the aggregate level, can reproduce a series of phenomena observed in the markets for cultural products. The models introduced in this thesis yield new insights in regard to the link between micro and macro-behavior. In addition, they can be seen as case-studies of interdisciplinary work bridging psychology and the social sciences. In what follows, I examine some of the general implications of the models at the applied, conceptual, and methodological level.

5.1 Integrating cognitive and agent-based models

Computational methods are already broadly used in simulating cognition and social systems. This methodological common ground can be capitalized to directly ground models of complex social systems on psychological models. The latter are the rudimentary building blocks for theories of collective behavior. Cognitive modeling on the one hand, and agent-based modeling on the other, allows for a seamless integration of the cognitive and social influence processes in contexts with repeated local or global social interaction. Rational choice models, by contrast, are based on strong rationality assumptions that restrict their potential to capture actual psychological processes. Further, naive statistical models can capture processes only indirectly by specifying proxy variables that relate to the underlying processes.

In chapters three and four, I pioneered this bottom-up approach by presenting models that detailed how people modify their behavior when they are exposed to information about the beliefs and choices of other individuals. In the opinion dynamics model presented in chapter three, this information was acquired locally by way of one-to-one interactions between the agents. In the popularity heuristic model presented in chapter four, information from individual decisions was, as in a bestseller list, added to a common informational pool to which all the decision makers had access. Both these models illustrated how psychological behaviors, such as the concurrent revision of opinion and confidence and popularity-driven search can lead to collective behavioral patterns such as opinion segregation or convergence in large groups and final outcome unpredictability in markets.

5.2 Psychological process models and prediction

Prediction is one of the main currencies by which the worth of models can be assessed. To start with, models can be compared to each other on the basis of the testable predictive hypotheses they generate. Further, models can be pitted against each other on the basis of the behavioral predictions they produce at the individual level.[1] Let us examine the predictive potential of the studies included in this dissertation.

First, the search and stopping theory introduced in chapter two can be directly compared to the probit framework largely in use in econometrics (see McFadden, 1980). Note that the environments postulated by our theory and by the probit model are essentially identical. In fact, when the cost of search equals zero and when there are only two the alternatives in the environment our approach converges to the probit model while for more than two alternatives it converges to the ordered probit model. Our model reveals that a non-negligible cost of search drastically increases the probability that alternatives in the first positions of the constructed order will be chosen while it radically decreases the chances that alternatives in lower positions in the order will be selected. Therefore, we expect that our ordered search models will clearly outperform probit and ordered probit models especially in real world choice environments with considerable costs of search. Moreover, our ordered search approach generates more behavioral predictions than the rational choice models it integrates. Our models predict the exact sequence in which alternatives should be examined and the eventual choice. One-shot multi-attribute utility models predict only the final choice. Optimal stopping theory predicts only the sequential stopping decisions. Further, in most scenarios it is impossible for optimal stopping models to predict the selected alternative ex-ante because alternatives

[1] Psychological process models produce rich behavioral predictions. For some interesting cases see Gigerenzer et al's (1991) probabilistic mental models and (2010) and Pleskac and Busemeyer's (2010) dynamic signal detection theory.

are sampled at random.

The opinion formation model presented in chapter three generates results about (i) the probability of convergence, polarization or fragmentation of opinions in large groups of people, (ii) the quality of the overall opinion before and after the convergence process (iii) the influence exerted by experts and naive majorities to the overall opinion of the group. Note that the model is one of the first to explicitly take into account the variable of relative confidence and to provide predictions in environments where individuals with different degrees of confidence influence each other. Thus, it can produce several new hypotheses that can be tested against data from the field and it can be applied in situations in which experts attempt to influence the public opinion.

Finally, the popularity heuristic model presented in chapter four generates several results about the markets of cultural products as a function of two variables — the diversity of taste in the population and the stopping thresholds characterizing the agents. The model can be used to estimate (i) the amount of market share inequality (ii) the outcome unpredictability and (iii) the imperfect correlation between quality and popularity. In addition, it reproduces the predictable failure but unpredictable success pattern characterizing these markets. Note that models assuming perfect knowledge can produce only end-state market-share distributions (result i). Further, models based on stochastic processes can dynamically reproduce the unpredictability and the end-state distributions (results i and ii), yet they do not take into account quality differences. By contrast, our framework can be used to derive hypotheses about any of these measures that could be tested in future large-scale studies such as in the music lab experiment. Moreover, at the purely behavioral level our model generates predictions about the exact search and stopping behavior of individual agents.

The projects included in this thesis demonstrate that psychological process models generate rich behavioral predictions at the individual level and produce comparatively more novel hypotheses about aggregate behavior than their as-if competitors. In the future, these predictions can guide the collection of data in the lab and in the field.

5.3 The value of simplicity and applicability

The theoretical analysis of chapter two revealed a surprising conceptual result. Two well-studied theories of rational choice, one-shot multi-attribute utility theory and optimal stopping theory, can be conceived as boundary cases of a generalized ordered search model in which individuals use a psychological estimation model to guide their search for good alternatives. In the past these two boundary cases have been studied independently of each other leading to very different conclusions about consumer behavior in markets. Moreover, in our framework the search and stopping problem has an intuitive optimal solution which can be expressed in terms of sequential cost-benefit analysis. Weitzman (1979), who was the first to work on this type of problem, discussed optimality within a dynamic programming framework. Weitzman's work is theoretically important, however, his algorithm cannot be readily conveyed or applied to real world scenarios. This might be one of the reasons why Roberts and Lattin (1991), who developed a model of consideration set formation based on the notion of ordered search, failed to notice the connections between their work and search models. By formalizing the ordered search problem in terms of plausible psychological models we made it sufficiently intuitive and we exposed the continuum between multi-attribute choice and optimal stopping models assuming random search. As process models are often easier to work and to communicate with they can become instruments for theory integration (for another example see Luan et al., 2011).

5.4 Correspondence rather than coherence

As the focus of psychological process models is correspondence to real world behavior rather than the internal coherence of the model, they spare us from committing type three errors, namely, finding the right answer to the wrong question. Statisticians and later economists have studied extensively optimal stopping problems with random search. Different versions of these problems have been explored by modifying some of the initial assumptions (e.g.

Rothschild, 1974; Conlisk, 2003). The ordered search models presented in chapter two and four indicate that in real world environments optimal stopping problems with random search are rarely encountered. Usually people have some additional information that they can use to guide their search. Thus, rather than sampling from a known utility distribution, people need to learn the relationship between the utility and the attributes. In fact, psychologists have already extensively investigated scenarios in which people learn linear or non-linear multivariate functions (e.g. Busemeyer et al., 1997). In the future, this knowledge can be directly transfered to the study of ordered search problems.

5.5 Concluding remarks

Psychology is the science that most diligently investigates the behavior of the most crucial factor in the study of social phenomena — the human agent. Therefore, it is the most appropriate discipline to ground and test the micro-foundations for the study of aggregate behavior. Precise theories are necessary in order to systematize the acquired psychological knowledge and to communicate it efficiently to neighboring disciplines. In this thesis I have presented three examples of psychological models that can be used to investigate problems of individual and aggregate behavior which have been explored so far in the social sciences (for more work in that direction see Mason et al., 2008; Goldstone and Gureckis, 2009; Denrell and Le Mens, 2013). I have argued that psychological models have several advantages over their as-if and statistical competitors. Such models produce richer predictive hypotheses, are more relevant to real world problems and are easier to convey and to work with. Undoubtedly, the future of the behavioral and social sciences will become increasingly interdisciplinary. This thesis signified another step in that direction and stressed the crucial role that psychology can play.

5.6 References

Busemeyer, J. R., Byun, E., Delosh, E. L., and McDaniel, M. A. (1997). Learning functional relations based on experience with input-output pairs by humans and artificial neural networks. In Lamberts, K. and Shanks, D., editors, *Knowledge, concepts and categories. Studies in cognition*. The MIT Press.

Conlisk, J. (2003). A note on convergence of adaptive satisficing to optimal stopping. *Journal of political economy*, 111(6):1353–1360.

Denrell, J. and Le Mens, G. (2013). Information sampling, conformity and collective mistaken beliefs. In *in M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), Proceedings of the 35th Annual Conference of the Cognitive Science Society.*, pages 2177–2182. Austin, TX: Cognitive Science Society.

Gigerenzer, G., Hoffrage, U., and Kleinbölting, H. (1991). Probabilistic mental models: a brunswikian theory of confidence. *Psychological review*, 98(4):506.

Goldstone, R. L. and Gureckis, T. M. (2009). Collective behavior. *Topics in Cognitive Science*, 1(3):412–438.

Luan, S., Schooler, L. J., and Gigerenzer, G. (2011). A signal-detection analysis of fast-and-frugal trees. *Psychological Review*, 118(2):316.

Mason, W. A., Jones, A., and Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, 137(3):422.

McFadden, D. (1980). Econometric models for probabilistic choice among products. *Journal of Business*, 53(3):13–29.

Pleskac, T. J. and Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological review*, 117(3):864.

Roberts, J. H. and Lattin, J. M. (1991). Development and testing of a model of consideration set composition. *Journal of Marketing Research*, 28:429–440.

Rothschild, M. (1974). Searching for the lowest price when the distribution of prices is unknown. *The Journal of Political Economy*, 82(4):689–711.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, 47(3):641–654.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig und ohne die unzulässige Hilfe Dritter verfasst habe und die Dissertation auch in Teilen keine Kopie anderer Arbeiten darstellt. Die verwendeten Hilfsmittel sowie Literatur sind vollständig angegeben. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden. Ich habe keinen Doktorgrad in dem Promotionsfach Psychologie, und die zugrunde liegende Promotionsordnung der Humboldt-Universität vom 03.08.2006 ist mir bekannt.

Berlin, den 03. Dezember 2014

____

Pantelis Pipergias Analytis