Understanding Complex Systems

Springer Complexity

Ragupathy Venkatachalam *Editor*

Artificial Intelligence, Learning and Computation in Economics and Finance

Springer

Springer Complexity

Springer Complexity is an interdisciplinary program publishing the best research and academic-level teaching on both fundamental and applied aspects of complex systems—cutting across all traditional disciplines of the natural and life sciences, engineering, economics, medicine, neuroscience, social and computer science.

Complex Systems are systems that comprise many interacting parts with the ability to generate a new quality of macroscopic collective behavior the manifestations of which are the spontaneous formation of distinctive temporal, spatial or functional structures. Models of such systems can be successfully mapped onto quite diverse "real-life" situations like the climate, the coherent emission of light from lasers, chemical reaction-diffusion systems, biological cellular networks, the dynamics of stock markets and of the internet, earthquake statistics and prediction, freeway traffic, the human brain, or the formation of opinions in social systems, to name just some of the popular applications.

Although their scope and methodologies overlap somewhat, one can distinguish the following main concepts and tools: self-organization, nonlinear dynamics, synergetics, turbulence, dynamical systems, catastrophes, instabilities, stochastic processes, chaos, graphs and networks, cellular automata, adaptive systems, genetic algorithms and computational intelligence.

The three major book publication platforms of the Springer Complexity program are the monograph series "Understanding Complex Systems" focusing on the various applications of complexity, the "Springer Series in Synergetics", which is devoted to the quantitative theoretical and methodological foundations, and the "Springer Briefs in Complexity" which are concise and topical working reports, case studies, surveys, essays and lecture notes of relevance to the field. In addition to the books in these two core series, the program also incorporates individual titles ranging from textbooks to major reference works.

Indexed by SCOPUS, INSPEC, zbMATH, SCImago.

Series Editors

Understanding Complex Systems

Ragupathy Venkatachalam
Editor

Artificial Intelligence, Learning and Computation in Economics and Finance

Springer

Editor
Ragupathy Venkatachalam
Institute of Management Studies
Goldsmiths, University of London
London, UK

ISSN 1860-0832 ISSN 1860-0840 (electronic)
Understanding Complex Systems
ISBN 978-3-031-15293-1 ISBN 978-3-031-15294-8 (eBook)
https://doi.org/10.1007/978-3-031-15294-8

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedicated to
Prof. Shu-Heng Chen

Acknowledgements

There are quite a few people who I would like to acknowledge for their help and support, without which this volume could not have been possible. From the time I first mooted the idea of having a festschrift for Shu-Heng Chen, all the way to the very end, I enjoyed the constant support of Vela Velupillai. I am grateful for his guidance, support, and, more importantly, for introducing Shu-Heng Chen to me. A special thanks to Connie Wang, who was instrumental for the successful completion of this project. Her unwavering support and kindness throughout deserve a special mention. I sincerely thank all the authors for their hard work and interesting contributions. Their enthusiasm and commitment to celebrate Shu-Heng's work has been a formidable source of inspiration for me, especially in light of what seemed like never-ending challenges and delays posed by the pandemic. Constantinos Repapis was extremely generous with his time and helped me during the initial stages of this project. I am grateful to Astrid Van den Bossche, Diviya Pant, Edgar Gutierrez, Sarath Jakka, Selda Kao, Sunil Kumar, Ye-Rong Du, and all the referees for their help and assistance in various ways. Finally, I would like to thank Hisako Niko, along with other members of the Springer team for their excellent support (and patience) through the making of this volume.

Contents

1 Computational Thinking in Economics and Finance: Introductory Remarks 1
Ragupathy Venkatachalam and Shu G. Wang

2 *Uncomputabilities*, Games, Axioms, Proofs and *Artificial Intelligence* 13
K. Vela Velupillai

3 Logic and Epistemology in Behavioral Economics 27
J. Barkley Rosser Jr.

4 Agent-Based Computational Economics: Overview and Brief History 41
Leigh Tesfatsion

5 Sequential Monte Carlo Squared for Agent-Based Models 59
Thomas Lux

6 Toward a General Model of Financial Markets 71
Nihad Aliyev and Xue-Zhong He

7 Sand Castles and Financial Systems: A Sandpile Metaphor 101
Francesco Luna

8 A Systematic Review of Investor Attention: Measurements, Implications, and Future Directions 121
Dehua Shen and Chen Wang

9 What is the Market? The Essential Teachings from an AI Market Experiment 141
Yuji Aruka

10 Market Power and the Hayek Hypothesis: An Experimental Investigation 161
Chung-Ching Tai and Bin-Tzong Chie

11 Counterfactual Thinking and Causal Mediation: An Application to Female Labour Force Participation in India 187
Sunil Mitra Kumar and Ying-Fang Kao

12 When Supply and Demand do NOT Meet: Sraffa's Critique of Economic Theory Restated 207
Stefano Zambelli

13 Visualizing the Roles of Frequent Terms in LTEs Following Two Economic Crisis Trigger Articles 247
Brian J. Kokensparger

14 Design and Performance Metrics for Autonomous Human-Machine Teams and Systems (A-HMT-S) 277
W. F. Lawless

15 Thirty-Five Years of Computational Economics 291
Robert Marks

16 Extending Herbert Simon's "Science of Design": the Role of Collaboration and Users in Development of Technically Advanced Systems 307
Ben Vermeulen and Andreas Pyka

Contributors

Aliyev Nihad University of Technology Sydney, Finance Department, UTS, Ultimo, Australia

Aruka Yuji Chuo University, Hachioji, Tokyo, Japan

Chie Bin-Tzong Department of Industrial Economics, Tamkang University, New Taipei City, Taiwan

He Xue-Zhong International Business School Suzhou (IBSS) Business Building (BS), South Campus, Xi'an Jiaotong-Liverpool University, Suzhou, People's Republic of China

Kao Ying-Fang Experimentation Team, Machine Learning and AI Division, Just Eat, London, UK

Kokensparger Brian J. Computer Science, Design and Journalism Department, Creighton University, Omaha, NE, USA

Kumar Sunil Mitra India Institute and Department of International Development, King's College London, London, UK

Lawless W. F. Professor of Math and Psychology, Paine College, Augusta, GA, USA

Luna Francesco Office of Innovation and Change, International Monetary Fund, Washington DC, NW, USA

Lux Thomas Department of Economics, University of Kiel, Kiel, Germany

Marks Robert Economics, University of New South Wales, Sydney, Australia

Pyka Andreas Institute of Economics, University of Hohenheim, Stuttgart, Germany

Rosser J. Barkley Jr. JMU College of Business, James Madison University, Harrisonburg, VA 22807, USA

Shen Dehua School of Finance, Nankai University, Tianjin, P.R. China

Tai Chung-Ching Southampton Business School, University of Southampton, Southampton, UK

Tesfatsion Leigh Research Professor and Professor Emerita of Economics, Courtesy Research Professor of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA

Velupillai K. Vela Solna, Sweden

Venkatachalam Ragupathy Institute of Management Studies, Goldsmiths, University of London, London, UK

Vermeulen Ben IQIB, Bad Neuenahr-Ahrweiler, Germany

Wang Chen College of Management and Economics, Tianjin University, Nankai District, Tianjin, P.R. China

Wang Shu G. Department of Economics (retired), National Chengchi University, Taipei, Taiwan

Zambelli Stefano Department of Economics and Management, University of Trento, Trento, Italy

Chapter 1
Computational Thinking in Economics and Finance: Introductory Remarks

Ragupathy Venkatachalam and Shu G. Wang

Abstract This chapter provides a brief overview of computational thinking in Economics and Finance. It explores the nexus between developments in Machine Learning, Artificial Intelligence and Economics. It sketches Shu-Heng Chen's contributions to the field and gives a panoramic view of the chapters included in this volume.

Keywords Computational economics · Artificial intelligence · Machine learning · Finance · Learning

Algorithmic thinking is gradually becoming an indispensable feature of Economics and Finance. Computational modes of thinking increasingly influence how we model individuals, organisations, market interactions and macroeconomic dynamics. Consider some topics that are traditionally of interest to economists and scholars of finance: decision-making by economic actors; mechanisms through which agents learn, adapt and thrive in uncertain, complex environments; the impact of economic decisions by individuals, firms and institutions on others in an interconnected system; emergence and diffusion of innovations; understanding volatility in financial markets; modelling aggregate, macroeconomic dynamics and associated pathologies (inflation, unemployment, inequality); and modes of exercising control through policy. All these issues are being studied using computational methods to varying degrees within the field.

Computational thinking in Economics and Finance can take diverse forms, ranging from straightforward applications of computational tools to analyse economic data, study properties of theoretical models (e.g., Computable General Equilibrium models), simulate large-scale Agent-Based Models (ABMs) and network models, perform numerical simulations (Chen et al. 2018; Tesfatsion and Judd 2006), inves-

R. Venkatachalam (✉)
Institute of Management Studies, Goldsmiths, University of London, London, UK
e-mail: r.venkatachalam@gold.ac.uk

S. G. Wang
Department of Economics (retired), National Chengchi University, Taipei, Taiwan

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_1

tigate game-theoretic scenarios; all the way to developing full-fledged algorithmic frameworks such as computable economics (Velupillai 2010).

However, the role of computation and simulation in Economics is by no means a recent phenomena (Velupillai and Zambelli 2015). The idea of logical machines and the use of computing metaphors—both analogue and digital—have a long and rich tradition in Economics. Varieties of analogue computing techniques have been used by Irving Fisher, A. W. H. Phillips, Richard M. Goodwin, Robert H. Strotz, Otto Smith, Arnold Tustin, Roy Allen and Oscar Lange, among others, often to study economic dynamics (Velupillai 2004). Digital computation too has a rich history in the discipline. The idea of simulation, and the access it offers to counterfactual alternative worlds, has been fruitful, at times indispensable, to study a variety of theoretical, empirical and policy issues. Rapid developments in information technology, together with impressive growth in the size, frequency and variety of data available to us in recent years, have further underscored the need and potential for developing appropriate computational tools for economic analysis.

For instance, ABMs are increasingly used to model complex interactions in Economics and Finance (Tesfatsion and Judd 2006; Chen 2017; Chen et al. 2012; LeBaron 2006). Algorithmic rationality and algorithmic game theory are at the frontiers of research topics today (Halpern and Pass 2015; Roughgarden 2016; Schmid et al. 2021). Computational methods and simulations are heavily relied upon to perform counterfactual thought experiments in ways that surpass the barriers posed by traditional mathematical and statistical tools (Wager and Athey 2018; Venkatachalam and Zambelli 2022). On the empirical front, Machine Learning (ML) and simulations have offered effective methods to analyse and learn from new types of data. These computational tools have permeated various sub-fields of Economics, Finance, and different schools of economic thought (Cogliano et al. 2021).

1.1 Artificial Intelligence and Machine Learning: The Future?

Despite the greatly exaggerated claims about its demise, Artificial Intelligence (AI) is back, yet again, to capture our imagination. This rejuvenated interest has spilled beyond the cloisters of academic communities to feature in a wide range of public discussions. Even if it remains to be seen whether some of its promises—such as achieving general-purpose intelligence—can be realised anytime soon, recent developments in AI seem to have persuaded scholars in a variety of disciplines to explore and employ computational methods in their research. This willingness to embrace computational thinking more widely has ushered fertile research programmes in many areas ranging from Biology, Epidemiology, Medicine, and Mathematics to Forensic Science, Law, Sociology, Transport Studies and more. Economics and Finance are no exception to this broader trend.

The question of what really constitutes AI continues to be debated. We prefer to take a broader, more plural view of AI to cover a broad umbrella of concepts, approaches and tools. These include: symbolic and sub-symbolic, connectionist approaches to AI, a range of machine learning methods such as supervised learning (e.g., artificial neural networks, nearest neighbour methods, kernel smoothening, probably approximately correct (PAC) learning, tree-based methods, classification algorithms, and support vector machines), unsupervised learning (e.g., self-organising maps, k-means, clustering), semi-supervised learning, ensemble learning (e.g., random forests), reinforcement learning, deep learning, natural language processing, evolutionary computation, ABMs, causal inference tools, ant algorithms and swarm intelligence.

It is worth noting that the dialogue between AI and Economics is not entirely new. This can be traced back to Herbert Simon's prolific work on Economics, AI and Cognitive Psychology (Simon and Newell 1972). In fact, Simon, along with Allen Newell, attended the famous Dartmouth workshop on AI in 1956, which many regard as the starting point of AI as a separate field.[1] Simon's approach to Economics was rooted in modelling individuals and organisations as problem-solvers with computational and cognitive limitations (bounded rationality), and analysing them through an algorithmic framework.

Why is AI relevant for Economics and Finance? A simple answer, in our view, would be the following: In so far as Economics and Finance focus on *how* human beings make decisions, deal with intractable and complex problems, or need to analyse big data sets, insights from AI and ML can prove to be important. Further, computational modes of theorising offer powerful ways to go beyond traditional formal methods.

This brings us to the next question: *which approach to AI is best suited for Economics?* Some key differences between the Classical and Modern approaches to AI, in terms of goals and focus, Russel et al. (2013), may be of some relevance here. If one is interested in utilising artificial agents to adequately encapsulate or learn about human behaviour, Classical AI continues to remain relevant. The Classical approach uses machines—digital computers in particular—as a vehicle to gain insights into *human* cognition, intelligence and learning. In this case, *explanation* of human behaviour, connection to actual cognitive limits and resemblance to human-like modes of reasoning are important (Simon 1983; Kao and Venkatachalam 2021). In contrast, the key goals for much of the Modern approach appear to be accurate prediction, and achieving human-level or superhuman performance in selected tasks and domains. Relatively less emphasis is laid on the degree to which these algorithms resemble *actual* human capabilities or behaviour in terms of decision-making or learning. This can certainly be fruitful in some areas, depending on the needs of the researcher. There are important differences between human learning and machine

[1] Herbert Simon is the only person so far to win both the Turing Award and the Nobel Memorial Prize in Economic Sciences. Allen Newell, Cliff Shaw, and Herbert Simon developed one of the earliest programmes—*Logic Theorist*—which utilised insights from problem solving skills of human beings, especially *heuristics*, to prove theorems in *Principia Mathematica*. The foundational paper of the field, however, was by Alan Turing (1950). Also, see Velupillai (2018).

learning, which in turn has implications for the choice of approach.[2] Recent advances in AI and ML have demonstrated impressive performance in board games like GO against top-level human players, relying on deep neural networks with reinforcement learning and Monte-Carlo tree search (Silver et al. 2016, 2017). Similarly, they have been highly successful in predicting protein structures with atomic-level accuracy (Jumper et al. 2021), and helped shed light on theoretical issues such as the mathematical structure of knots or hidden symmetries (Davies et al. 2021). In the case of the latter, AI tools help develop intuition concerning mathematical challenges, act as a heuristic aid or guide, and help develop collaborative approaches to solve difficult problems.

There is reason to hope that such breakthroughs are also possible in Economics and Finance. These advances in AI, ML, ABMs, complex systems and big data, together have offered new ways of thinking coherently about economic problems in an algorithmic mode, thus opening promising avenues for research (Mullainathan and Spiess 2017; Athey 2019; Agrawal et al. 2019; Mullainathan and Obermeyer 2022). To focus solely on ML and developments along the lines of statistical learning theory would be to miss valuable insights. For instance, while ML plays an important role in estimation, there are important questions of identification and discovering causal mechanisms that often require broader approaches. Equally, to entirely forgo the power and possibility offered by modern AI and ML would be unwise. The actual choice should be dictated by the nature of the problem being studied; a blend of the two could offer richer possibilities.[3] Developments in causal inference promise an important link between these two approaches. As computational approaches become crucial to the social sciences, they also raise important questions concerning methodology, epistemology, issues related to prediction, validation, and inference. In our view, these questions will continue to occupy scholars in the future.

1.2 Shu-Heng Chen and Computational Social Sciences

This volume is a festschrift in honour of Prof. Shu-Heng Chen, who has made several pioneering contributions to Computational Economics, Behavioural Economics and Finance, AI, Learning, ABMs, Econophysics, Evolutionary Computation and other

---

[2] It is worth distinguishing here between concerns that relate to *efficiency* and *effectivity*. As (Simon, 1983, p. 27) points out:

> Artificial intelligence has two goals. First, AI is directed toward getting computers to be smart and do smart things so that human beings don't have to do them. And second, AI (sometimes called cognitive simulation, or information processing psychology) is also directed at using computers to simulate human beings, so that we can find out how humans work and perhaps can help them to be a little better in their work.

[3] There is now a growing literature on the impact of AI on various aspects of the economy and society, see Agrawal et al. (2019).

related areas. He is an important contributor to the development of Computational Economics, and has been among the early advocates of using AI tools to understand economic problems, dating back to the 1990s. The festschrift was originally planned for his 60th birthday, but was delayed in the event of the several challenges posed by the COVID-19 pandemic.

Shu-Heng was born in 1959, in Taizhong, Taiwan. He joined the National Chengchi University (NCCU) in 1977 to pursue his undergraduate studies in Economics. He then moved to the prestigious National Taiwan University (NTU) in 1983 to continue his Masters degree and subsequently to the University of California, Los Angeles (UCLA) in 1984 to embark on a PhD in Economics. At UCLA, Shu-Heng's horizons broadened beyond Economics as he was exposed to a range of subjects and developed a particular interest in computational methods. Vela Velupillai and John McCall, his mentors at UCLA, have had a lasting intellectual influence on him. Shu-Heng was a part of the Centre for Computable Economics at UCLA and attended lectures by Velupillai on topics related to computability theory. His doctoral thesis, titled '*On the Complexity in Adaptive Economic Systems: the relation between RBS and PDP in Adaptive Economic Systems*' dealt with issues concerning complexity and had applications for the Taiwanese stock market.[4] Upon graduating in 1992, Shu-Heng returned to Taiwan, to his *alma mater*, NCCU to embark on what was to become a highly prolific and successful career and a wonderful adventure. Shu-Heng is currently a Distinguished Professor in the department of Economics at NCCU. In 1995, Shu-Heng started AI-ECON, one of the first research centres in the world dedicated to exploring the nexus between AI and Economics, thus developing a systematic research programme. The intellectual seeds for AI-ECON had already been planted during his time at UCLA. In particular, the works of John von Neumann and Arthur W. Burks on the theory of self-reproducing automata (von Neumann 1966); contributions to genetic algorithms by John Holland, to Genetic Programming by John Koza; and work related to Economics and Complex Systems done at the Santa Fe Institute, prominently influenced the direction of research undertaken by Shu-Heng in the years that followed.

Shu-Heng is probably best known for introducing Genetic Programming and other evolutionary algorithms into Economics, and for his role in the development of Agent-Based Computational Economics (Chen and Yeh 1997, 2001). His initial years as a young academic were focused on developing this biologically-motivated approach to understanding the evolution of economic systems, markets, and other institutions. Shu-Heng saw this as a fruitful way to emancipate Economics from the shackles of a static *homo economicus* framework. He was in search of tools to encapsulate what he saw intuitively as important characteristics of the economic system: the evolutionary nature of the system, agents with bounded rationality, and the importance of interaction, adaptation and learning. To this end, he adopted, developed, and adapted a computational approach to further his research pursuits over the years.

The connections between Shu-Heng's work and contributions by Simon in this context are quite easy to understand (Chen and Kao 2016). Like Simon, Shu-Heng

[4] Velupillai's contribution to this volume provides more details about Shu's time at UCLA.

was interested not just in developing computational tools, but in using them effectively to understand *human* behaviour. In this regard, Shu-Heng developed an innovative research programme to combine the knowledge gained from Experimental Economics regarding cognitive constraints with the power of computational methods. His approach centred around using knowledge from the former to design more *human-like* agents in the latter, trying to bridge the gap between simulated and actual worlds. His contributions to experimental prediction markets, social networks and trust, for instance, are examples of this broader approach. Shu-Heng has made remarkable contributions to the field of Finance, which has been a field of interest since the beginning of his career (Chen and Yeh 2002; Chen et al. 2001). His co-edited volume '*Computational Intelligence in Economics and Finance*' (Chen and Wang 2004) bears testimony to his pioneering vision, in which many of the AI and ML algorithms popular today had already been broached, and their applications cogently outlined by him over two decades ago. This work has been continually evolving and has matured into a coherent computational approach to Behavioural Economics (Chen et al. 2016).

Shu-Heng has become more interested in the methodological and philosophical questions in Economics (Chen 2012, 2020; Chen and Venkatachalam 2017) during the later part of his career, and the field of Digital Humanities has been fortunate to capture his attention in the recent past. Shu-Heng's openness to interdisciplinary approaches and his legendary work ethic are well-known among those of us who have had the opportunity to work with him or to learn from him.[5] In addition to the various awards, books, over three-dozen edited volumes, hundreds of papers, book chapters, conference proceedings and editorial positions, Shu-Heng has also shown admirable dedication to institution-building by serving in various senior administrative roles in NCCU.

A particularly noteworthy personal characteristic of Shu-Heng has been his support to young scholars. Life was tough for Shu-Heng in his early years at NCCU. At one point, he even had to take adjunct teaching positions in other universities in order to earn extra money to support his students. In 2000, when the School was relocated to a more spacious building, Shu-Heng took advantage of the *rental* opportunities released from the Dean so that his students were among the very lucky few graduate students in NCCU to have their own research spaces: a plain desk, instead of a luxurious room, per person. This marked the embryonic form of the AI-Econ Research Center, which he later formalised as his research team. Since then, the AI-ECON Research Center has become a welcoming and accommodating hub for very many talented people, including students, post-doctoral researchers and visiting scholars. The Center has never been a part of the School's organization chart and therefore does not receive a regular budget from the University. Every penny poured into the Center for the support of its members and activities (such as conducting research projects and hosting conferences) came from Shu-Heng's own salaries, research grants, goal-specific funding, constant bullion-sized donations from himself and more rarely a-penny-here-a-nickel-there donations from a small number

[5] To this list, we would also add his love for spicy noodles!.

of supportive colleagues. It remains as a *pay-as-you-go* research center under the School of Social Sciences at NCCU up to this day, despite the fact that it has long been an unmissable attraction of the School.[6]

Discussing Shu-Heng's career and achievements wouldn't be complete without mentioning an important member of his team. As his career and the Center grew, *logistics* and *management* issues posed a significant burden, which fell upon Shu-Heng. He was extremely fortunate to have had the able and generous support of Ms. (and later Dr.) *Connie Houning Wang*, who joined him at a critical moment. Connie happens to be an old acquaintance of Shu-Heng since they were both students at the NCCU. Her incredible administrative skills have been crucial for developing Shu-Heng's research programme and her contribution to the Center can never be overestimated. Connie was able to put every penny to most effective use. In addition to administering the Center, she helped organise academic conferences and countless meetings, and oversaw the research progress of the members in the Center.

1.3 Brief Overview of the Contributions

This volume brings together an admirable range of contributions related to AI, learning and computation—themes that remain the foundational pillars of Shu-Heng's research in Economics and Finance. The chapters touch upon one or more of these themes in varying degrees and from different viewpoints, thus painting a rich collage. The contributors are predominantly Shu-Heng's collaborators, students, friends and admirers from many different locations, subject areas and generations—all of which speaks to the wide reach of his work.

The two chapters that immediately follow this introduction deal with topics related to logical and mathematical foundations of computation. Velupillai (Chap. 2) deals with deep and foundational questions concerning computability, logic, and proofs. He invokes the metaphorical duel between Hydra and Hercules, and links it imaginatively to a finite, arithmetical game that discusses a variety of issues concerning computability, intuition, and axiomatics. This has direct relevance for attempts to link human thinking to machine learning, via AI. Rosser's chapter (Chap. 3) explores epistemological issues in the context of complex dynamical systems. He distinguishes between dynamic and computational complexity, both of which offer interesting insights. The possibilities offered by Simon's approach to Behavioural Economics that address some of these problems are also explored. Rosser Jr. discusses the foundations of Simon's concept of 'bounded rationality', which is motivated by limits to computation by human actors, and elaborates its connection to coping with complexity and limits of knowledge.

[6] Especially, whenever the Deans need to take international visitors around the School for a guided tour!.

The next two chapters deal with ABMs in Economics and Finance. Tesfatsion (Chap. 4) provides an excellent overview, recounting the history and development of Agent-Based Computational Economics (ACE). It offers us a valuable account of the field through the eyes of someone who pioneered and contributed immensely to it, along with Shu-Heng and other scholars. ACE is certainly an important framework that has gained a lot of popularity over the years. It offers a lot of potential and flexibility for computational experiments, especially in light of the challenges that interactions, coupling, feedback effects and complex dynamics can pose for traditional modelling techniques. Tesfatsion outlines the idea of completely Agent-Based Modeling (c-ABM), which refers to 'the computational modeling of processes as open-ended dynamic systems of interacting agents' as characterised by seven modelling principles.

Lux (Chap. 5) takes up the question of whether there are alternatives to standard Markov Chain Monte Carlo (MCMC) algorithms that exist, which can make Bayesian estimation possible and easier to implement in the context of ABMs. The chapter proposes a combination of two Sequential Monte Carlo algorithms (one each for the hidden states of the system, and the parameters) for this purpose. Lux uses the model of Alfarano et al. (2008) as a test case for the estimation, and shows that the computational demands are much lower compared to the standard MCMC methods for Bayesian estimation.

This is followed by four chapters that deal with a range of issues related to financial markets. Aliyev and He (Chap. 6) develop a novel framework to understand financial markets by combining insights from Decision Theory and Information Science. The framework lays focus on the precision and reliability of information as well as adopting a broad definition of subjective rationality. In the framework, there are multiple decision theories corresponding to each class of information, and they show how behavioural anomalies at the individual and market level can be better understood. Luna's chapter (Chap. 7) deals with the issue of complexity in financial markets. It explores whether there is a level of interconnection among financial institutions that will maximize the welfare of the system by minimizing the average loss per institution, per period of time. Using the metaphor of the sand pile, he reformulates this question to assess the effect of placing the falling sand grains (financial shocks) to the location in the pile that is less likely to collapse. Through simulations, the chapter explores the intuition that pooling financial resources is likely to reduce the number of financial distress episodes for individual institutions, yet at the same time making the fewer crises significantly more severe for the system as a whole. Shen and Wang (Chap. 8) explore the role of investor attention in financial markets in a detailed manner. Limits to attention, a scarce cognitive resource along with limits on human information processing abilities, have been important themes in Simon's research programme. Shen and Wang survey the literature in this area, explain why and how investor attention matters, and its influence on behaviour and financial market dynamics.They also investigate issues related to the measurement of investor attention. This provides useful ingredients for designing artificial agents, so as to faithfully capture important characteristics of human participants in financial markets.

The next three chapters address learning and inference through experiments and counterfactual reasoning, along with computational methods. Aruka (Chap. 9) examines the settlement mechanism of the Bitcoin Exchange and compares it with the stock exchange. The chapter utilises the idea of digit length frequency distribution in the context of Fully Random Iterated Cellular Automata (FRICA). This idea is applied to the price time series generated by the U-Mart acceleration experiment tool. Using this, the chapter examines the neutralizing effects due to the change in market transaction strategy composition. Tai and Chie (Chap. 10) take up the *Hayek hypothesis*, which posits that market prices will converge to a competitive equilibrium as agents adapt and learn over time. They examine any observed violations of this hypothesis in their experiments. Unlike other studies, Tai and Chie choose to examine this from a cognitive viewpoint. The chapter asks whether the cognitive ability of participants has any observable influence on behaviour in (experimental) auction markets, and the extent to which this can be used to explain market dynamics. In their experiments, each human participant plays against a group of artificial (robot) agents. They do not find evidence that higher cognitive scores of players are associated with better price convergence. Kumar and Kao (Chap. 11) deal with one of the frontier topics in ML and AI, viz. causal inference. While ML techniques have proven to be extremely powerful in prediction and classification tasks, social scientists are often concerned with *explanations* and causal *mechanisms* that are at play. The need to move beyond ML's focus on goodness of fit and prediction to a paradigm where seeking causal explanations are a priority, has in turn led to a flurry of work in the interface between AI, ML, and causal inference. This has prompted computational social scientists to think about ways in which learning about causal effects can complement priorities concerning predicting outcomes. Some ways to combine them through integrative modelling are outlined in Hofman et al. (2021). Kumar and Kao demonstrate how causal models can be effectively used to analyse observational data and answer useful counterfactual queries, building on the work by Judea Pearl and others on Directed Acyclic Graphs (DAGs), and identifying path-specific effects. They apply a causal mediation framework to study the effects of education on female labour-force participation in the Indian context.

The wide-ranging applications of computational methods to theory, practise, and design constitute the shared theme of the following chapters. Zambelli (Chap. 12) revisits the well-known critique presented by Piero Sraffa in the book Production of Commodities by Means of Commodities (Sraffa 1960) and reinvestigates it using rigorous algorithmic methods. The essence of the critique, he argues, 'is that standard economic theories do not contain enough ingredients to demonstrate that book-keeping *equilibrium prices* are determined by the working of *free* or *perfect* competition', which in turn implies that one cannot infer that market forces alone determine distribution. He presents a detailed, admirable algorithmic exploration of Sraffian schemes through numerical simulations and concludes that the critique still stands. This exercise also illuminates how computational tools can be useful to investigate important questions from different schools of economic thought. Lawless (Chap. 13) looks at how metrics for the design and performance of a team or an organization need to be transformed to assist autonomous systems with making decisions in uncertain environments. The chapter proposes a theory of the interdependence of complementarity. This is particularly relevant for developments in AI that focus on autonomous Human-Machine Teams. They also have relevance for Economics, where decision-making takes place in noisy, incomplete information environments.

Kokensparger (Chap. 14) examines the Letters to the Editor (LTEs) written for a newspaper over a twelve-week period, after the publication of articles reporting the likelihood of sustained impact of two oncoming economic crises (the oil crisis and the COVID pandemic). The chapter employs computational techniques to analyse the frequency of terms and the shapes of responses, thereby providing novel ways to think about narratives that people form in an economy. Textual analysis through computation is gaining attention among researchers within Economics and Finance, thus expanding the traditional conceptions about data to analyse public sentiments. Marks (Chap. 15) describes his multi-faceted work in the area of computational economics and recounts his research programme that spans over 30 years. It starts from his development of algorithms to play a generalisation of the Iterated Prisoner's Dilemma (IPD), applying a then-new method of machine learning (genetic algorithm, GA) to this problem, to extending some of these techniques to oligopoly pricing. He later traces how he developed ABMs, applying GA to exploring the best method of decision making in uncertain situations, and his use of simulations to search exhaustively among methods for decision-making under uncertainty. In the final chapter (Chap. 16), Vermeulen and Pyka revisit Simon's framework concerning the 'science of design' to introduce a collaboration of developers as well as users in the invention, development, and design of an advanced system (e.g., an aircraft). They extend Simon's theoretical concept of iterative tree traversal for technical system development, highlighting 'the co-evolutionary interaction of the system development process with search for the organization of the project with various experts and market exploration'. Vermeulen and Pyka outline a stylistic model of the process of collaborative technology development. This offers several interesting ideas for thinking about design and the practical aspects of dealing with market and technological uncertainty from a development point of view.

It is our hope that these chapters give readers a taste of the latest developments in the field, and help spawn interesting ideas for future interdisciplinary research. The direction offered by computational methods, at least to the extent that can be glimpsed ahead, shows much promise to enrich and advance our understanding of economic systems. It is a road that we believe is worth travelling.

References

Agrawal A, Gans J, Goldfarb A (eds) (2019) The economics of artificial intelligence: an agenda. University of Chicago Press, Chicago

Athey S (2019) The impact of machine learning on economics. In: Agrawal A, Gans J, Goldfarb A (eds) The economics of artificial intelligence: an agenda. University of Chicago Press, Chicago, pp 507–547

Chen S-H (2012) Varieties of agents in agent-based computational economics: a historical and an interdisciplinary perspective. J Econ Dyn Control 36(1):1–25
Chen S-H (2017) Agent-based computational economics: how the idea originated and where it is going. Routledge, Abingdon
Chen S-H (2020) On the ontological turn in economics: the promises of agent-based computational economics. Philos Soc Sci 50(3):238–259
Chen S-H, Chang C-L, Du Y-R (2012) Agent-based economic models and econometrics. Knowl Eng Rev 27(2):187–219
Chen S-H, Kaboudan M, Du Y-R (eds) (2018) The Oxford handbook of computational economics and finance. Oxford University Press, Oxford
Chen S-H, Kao Y-F (2016) Herbert Simon and agent-based computational economics. In: Frantz R, Marsh L (eds) Minds, models and Milieux: commemorating the centennial of the birth of Herbert Simon. Palgrave Macmillan, London, pp 113–144
Chen S-H, Kao Y-F, Venkatachalam R (2016) Computational behavioral economics. In: Frantz R, Chen S-H, Dopfer K, Heukelom F, Mousavi S (eds) Routledge handbook of behavioral economics. Routledge, Abingdon, pp 309–331
Chen S-H, Lux T, Marchesi M (2001) Testing for non-linear structure in an artificial financial market. J Econ Behav Organ 46(3):327–342
Chen S-H, Venkatachalam R (2017) Agent-based models and their development through the lens of networks. In: Aruka Y, Kirman A (eds) Economic foundations for social complexity science. Springer, Singapore, pp 89–106
Chen S-H, Wang PP (2004) Computational intelligence in economics and finance. In: Chen S-H, Wang PP (eds) Computational intelligence in economics and finance. Springer, pp 3–55
Chen S-H, Yeh C-H (1997) Toward a computable approach to the efficient market hypothesis: an application of genetic programming. J Econ Dyn Control 21(6):1043–1063
Chen S-H, Yeh C-H (2001) Evolving traders and the business school with genetic programming: a new architecture of the agent-based artificial stock market. J Econ Dyn Control 25(3–4):363–393
Chen S-H, Yeh C-H (2002) On the emergent properties of artificial stock markets: the efficient market hypothesis and the rational expectations hypothesis. J Econ Behav Organ 49(2):217–239
Cogliano JF, Veneziani R, Yoshihara N (2021) Computational methods and Classical-Marxian economics. J Econ Surv 36(2):310–349
Davies A, Veličković P, Buesing L, Blackwell S, Zheng D, Tomašev N, Tanburn R, Battaglia P, Blundell C, Juhász A et al (2021) Advancing mathematics by guiding human intuition with AI. Nature 600(7887):70–74
Halpern JY, Pass R (2015) Algorithmic rationality: game theory with costly computation. J Econ Theory 156:246–268
Hofman JM, Watts DJ, Athey S, Garip F, Griffiths TL, Kleinberg J, Margetts H, Mullainathan S, Salganik MJ, Vazire S et al (2021) Integrating explanation and prediction in computational social science. Nature 595(7866):181–188
Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A et al (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596(7873):583–589
Kao Y-F, Venkatachalam R (2021) Human and machine learning. Comput Econ 57(3):889–909
LeBaron B (2006) Agent-based computational finance. In: Tesfatsion L, Judd K (eds) Handbook of computational economics 2: agent-based computational economics, vol 2. Handbook of computational economics. North-Holland (Elsevier), Amsterdam, pp 1187–1233
Mullainathan S, Obermeyer Z (2022) Diagnosing physician error: a machine learning approach to low-value health care. Q J Econ 137(2):679–727
Mullainathan S, Spiess J (2017) Machine learning: an applied econometric approach. J Econ Perspect 31(2):87–106
Roughgarden T (2016) Twenty lectures on algorithmic game theory. Cambridge University Press
Russel S, Norvig P et al. (2013) Artificial intelligence: a modern approach. Pearson Education Limited, London

Schmid M, Moravcik M, Burch N, Kadlec R, Davidson J, Waugh K, Bard N, Timbers F, Lanctot M, Holland Z et al. (2021) Player of games. arXiv preprint arXiv:2112.03178
Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M et al (2016) Mastering the game of Go with deep neural networks and tree search. Nature 529(7587):484–489
Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A et al (2017) Mastering the game of GO without human knowledge. Nature 550(7676):354–359
Simon HA (1983) Why should machines learn? In: Michalskl R, Carbonell J, Mitchell T (eds) Machine learning - an artificial approach. Tioga Publishing Co, Palo Alto, pp 25–37
Simon HA, Newell A (1972) Human problem solving. Prentice Hall
Sraffa P (1960) Production of commodities by means of commodities. Cambridge University Press, Cambridge
Tesfatsion L, Judd KL (eds) (2006) Handbook of computational economics 2: agent-based computational economics. North-Holland (Elsevier), Amsterdam
Turing AM (1950) Computing machinery and intelligence. Mind: A Quart Rev Psychol Philos 59(236):433–460
Velupillai K (2010) Computable foundations for economics. Routledge, Abingdon
Velupillai KV (2004) Economic dynamics and computation - resurrecting the Icarus tradition. Metroeconomica 55(2–3):239–264
Velupillai KV (2018) Models of Simon. Routledge, Abingdon
Velupillai KV, Zambelli S (2015) Simulation, computation and dynamics in economics. J Econ Methodol 22(1):1–27
Venkatachalam R, Zambelli S (2022) Self-replacing prices with credit and debt. Struct Chang Econ Dyn 60:451–466
von Neumann J (1966) Theory of self-reproducing automata. University of Illinois Press, Urbana
Wager S, Athey S (2018) Estimation and inference of heterogeneous treatment effects using random forests. J Am Stat Assoc 113(523):1228–1242

Chapter 2
Uncomputabilities, Games, Axioms, Proofs and *Artificial Intelligence*

K. Vela Velupillai

Abstract The prevalence of *Undecidabilities*, *Incompleteness* and *Uncomputabilities* in formal systems do not make it easy to devise policies, especially proving their existence, even in *finite* problem spaces. The interplay between the *finite* and *infinite*, between the *logically formal* and the *intuitively sensible*, is difficult to disentangle. This is especially evident in computable and uncomputable games. In this paper these games and some foundational questions of a mathematical and logical nature are also explored.

Keywords Goodstein sequence · Growth of functions · Paris-Harrington theorem · Hercules-Hydra game · Arithmetic games · Busy Beaver games

2.1 Introductory Notes

> I probably met you [for the first time] on your 43nd birthday. [In] the fall semester [of 1990] at UCLA you opened a course on [Theories of the] Business Cycle, and my enrollment [in] your class changed [my] life, which can be christened as 'A course changes my life'.
>
> **Shu-Heng Chen**, e-mail to me on 25/9/2019.

My life, especially its intellectual component, was changed irrevocably by Shu-Heng[1] becoming my doctoral student from October, 1990. His outstanding UCLA doctoral thesis, conceived, written and defended within three years, was titled *On*

[1] I shall, henceforth, refer to Shu-Heng Chen as simply *Shu*.

'This is a finite constructive proof of the restricted ordinal theorem for ordinals less than ω^{ω}' Goodstein, 1944♣, p. 36; italics in the original.

♣Many of the articles cited in this paper are from a 'bygone' era, when gender neutrality was not even considered worthy of relevance.

K. V. Velupillai (✉)
Tottvägen 11, 169 54 Solna, Sweden
e-mail: kvelupillai@gmail.com

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_2

the Complexity in Adaptive Economic Systems: *the relation between RBS* [rule-based system] *and PDP* [parallel distributed processing] *in Adaptive Economic Systems*.

RBS has a long and respectable history in mechanical calculations—or example in the well-known case of long-multiplication (and division)—and has a clear machine-based or algorithmic interpretation. Machines that implemented *PDP* systems were relatively new. It was, thus, an interesting exercise to investigate the relative merits of the two methods in the context of adaptive economic systems and their complexity––whether computational, informational or Kolmogorov. Shu's investigations, with solid theoretical foundations, were of the actual data coming out of the Taiwanese stock market.

Partly as a result of his extreme competence, especially on a comprehensive, yet constructively critical study of the burgeoning literature on *PDP*, and partly because he was a naturally versatile (mathematical) programmer, Shu became an expert also on *connectionism*. He was able to translate *connectionism* in *computable* frameworks and investigate its universality properties depending on the kind of *neural* (or *brain*) structure that would act as a repository of *PDP*.

This obviously led to an attempted understanding of the elements of neuroscience and learning structures, mainly by me, and thus a reading of Cajal (1937; 1989), Sherrington (1906), Turing (1936, 1937), McCulloch and Pitts (1943), Rosenbleuth, Wiener and Bigelow (1943, Ashby (1945), Hebb (1949), Chomsky (1957), Gold's *Rand Research Memorandum* of 1964 (later Gold 1964)—but also to the 'modern' literature on PDPs, exemplified by the two volumes of Rumelhart et al., (1986a, 1986b) and McCall (1989). These had positive external effects on my intellectual development and contributed to the reorganization of my understanding of the nature and scope of what I came to call *computable economics*.

I began this fascinating journey, encouraged by Shu and John McCall, with the beautiful *Gifford Lectures* by Sherrington (1940), which led me to the conviction that there were many confluences, inspired—either explicitly or implicitly—by Ramon y Cajal at the turn of the 19^{th} to the twentieth century. It led to the cybernetic movement, for which the implicit inspiration came via Turing (*op.cit.*) and McCulloch and Pitts (*op.cit.*) and explicitly on the basis of Wiener (1948, 1961) and by Turing (1948; 1992).[2] Much later, after Shu had left the 'orbit' of my (intellectual) life, I came to understand that this 1940s confluence came to be organized as *The Ratio Club*—which was, in effect, the British *cybernetic group* (Husbands and Holland 2008, p. 91).

By the time Shu entered my (intellectual) life, I had begun to walk on the two feet of *computable economics* and *nonlinear dynamic processes*. An aspect of the computable 'leg' was represented by Shu's thesis; the other foot had been in motion, literally and metaphorically, ever since I had been a student of Thalberg (in Lund)

[2] And his thoughts and research that was to lead to the literature which he always called *Machine Intelligence*, later—unfortunately—referred to as *Artificial Intelligence*, a name conjured up in opposition to *cybernetics*. I read Turing's 1948 piece in an earlier published version, but for simplicity the reference here is to the later, easily available, publication.

and Goodwin (in Cambridge) and saw its apex in Zambelli's thesis and my courses at various universities on *Theories of Business Cycles.*

Soon after Shu left UCLA with his PhD, he arranged for me to visit National Chengchi University (NCCU) in the academic year 1992–93, where I met, also, Professors Mao Tai and Shu G. Wang. They, all three of them, impressed upon me the teachings of Confucius and Lao Tzu and the importance of walking on the two feet of theory and its applications (particularly in economics which they, on the basis of thousands of years of Chinese philosophy, viewed as advocating policy proposals, based on theoretical constructions).

It has taken over twenty years for their wise outlooks and teachings to take *some* root in my work—which, until then, concentrated solely on the theoretical sides of computable economics and nonlinear dynamics.

The two feet on which I attempt to walk are *based on* an applied part which is *supposed* to help in the modelling of the search and construction of a vaccine against the coronavirus epidemic. No vaccine is entirely efficacious—or 'full proof' against any virus. However, the main thrust of the applied part is to divest proofs of axioms and let them be determined by the conjunction of inference rules and computation rules. This must be borne in mind when the theoretical part, necessarily 'caricaturing' the chemistry, botany, zoology and many other factors of life-scientific and social-scientific nature. Proof, in my senses, also plays a crucial role in the sciences.

In the particular case of my *theoretical* foundations, they harness *logical* and *mathematical* aspects, but modulated by emphasizing the role of *human intuitions* in these formulations. They are based on re-interpreting the mythical dual between *Hydra* and *Hercules*[3] as *effectivised* arithmetical and *Busy Beaver* games—thus linking the computable and uncomputable.

The paper is organized as follows. The next section is primarily an explanation of the *Hercules* vs. *Hydra* game. The intuitive result of considering the game as an *effectivized* play of a restricted two-person scheme allows me to use a result of Rabin (1957). Section 3 is an interpretation of the *Hydra* vs. *Hercules* game in terms of Rado's *Busy Beaver game* (Rado 1962). In Sect. 4, there are some (skeptical) notes on the use of axioms in proofs, (of theorems) in mathematics and logic, in *any* formal sense. The final Sect. 5 is a kind of speculative conclusion on being able to walk on two feet, firmly planted on relatively immovable terrain.

---

[3] Hodgson (2004, p. 12; italics added) explains, in a paper that outlines, in a comprehensible way, many of the 'difficult' issues that are of relevance to this paper:

'The second of the twelve labours imposed on *Hercules* was a fight against the Lernæan *Hydra*, with [many] heads … which would grow again as soon as chopped off ….'.

Fig. 2.1 Shu G. Wang, Shu-Heng Chen and Velupillai in the guest house room at NCCU during the latter's visit to Taipei in the academic year, 1992–93

2.2 The Hydra-Hercules Battle as an Uncomputable Arithmetic Game

> The *Hydra* was a monster with many heads, living in a marsh of Lerna. *Hercules*[4] fought it in a formidable struggle, *chopping off its heads* with a club. However, as *Hercules* chopped ahead , two new heads grew forth from the monster's body. According to the official records *Hercules won.* What the Greek did not reveal is that, *no matter in which order Hercules had decided to attack the heads*, he would have annihilated the *Hydra.*
>
> Luccio and Pagli (2000), p. 130; italics added.

The Hydra-Hercules battle *can* be interpreted as a two-person, full-information, alternating, game, which terminates in finite time, provided the Greeks did reveal that Hercules always 'annihilated the Hydra'—whatever the order in which the mythological hero chooses to decapitate the monster. There is no 'customary' chance element since the battle—the game—starts only when Hercules chops off one of the heads of the Hydra! Thus, this is an arithmetical game; however, I shall *not* assume the axiom of determinacy. Therefore, the game may be undetermined. I shall look for a way to *effectively* solve this game (or puzzle, see below)—i.e., show its determinacy.

An interpretation of the battle, usually with pictures of a branching tree, with leaves and a root, is customary; this literature is replete with stylized figures of trees sprouting branches depicting Hydra's reaction to Hercules, wielding axe-like chopping devices. Hydra sprouts new branches, with leaves at the end of them—the former depicting new bodies, with the leaves as new heads, replacing the ones

[4] Just as I assume Hercules is a *male*, Hydra is an '*it*'—but see also the lead footnote.

chopped off by Hercules. But there are also other respectable way of analyzing this game—as a task which one completes, whether it is the objective to finish it, or not; or whether eggs, balls, etc., are packed into a box or emptied from it, and so on (Gardner 1983, Smullyan 1979).

This standard literature on the resolving of the outcome of the battle—the *termination*[5] of an *algorithm* or the convergence of a sequence—takes off from Goodstein (1944), considering the pioneering article by Paris and Harrington (1977) to be a 'natural' extension, exhibiting the limitations of the *Peano Axioms* (PA, for the arithmetic of natural numbers) for the *proof* of 'termination' of the relevant sequence, and the inevitable need to go beyond *to* the infinities of Cantor's ordinal numbers. The celebrated Paris-Harrington result is supposed to be a 'natural mathematical theorem,'[6] as distinct from Gödel (1931) on incompleteness, which is supposed to be a logical construct, largely artificial. This results in the Kirby and Paris (1982) contribution to considering the battle between Hydra and Hercules with the full paraphernalia of ordinal theories, Ramsey growth of numbers and function, etc., in terms of the aforementioned tree paradigm. The trees, best interpreted graph-theoretically, the chopping-off of heads as the 'cutting of leaves,' the rules of replication in terms of the language of trees, has become standard and any number of the papers in this genre can be cited for figurative elucidation—but I confine my citations to two classics Lucio and Pagli (*op.cit.*) and Kirby and Paris (*op.cit.*)

Let me return to the arithmetical game interpretation of the battle between Hydra and Hercules and observe the following[7]:

- Hercules starts—i.e., makes the initial move in the game—by chopping off a head (leaf at the top of a branch of a tree with a root);
- There are no chance moves—as strategies in a game—in this battle between the Hydra and Hercules;
- The Hydra grows another head, by any (integer) multiple of the head chopped off, at the point in the branch of the tree where the leaf that was cut off by Hercules, in the previous stage;
- The chopped off head—cut off leaf, at the tip of a branch of a rooted tree—and the growing head(s) can be *coded* (as natural numbers);
- Hercules, in *finite time*, reaches the root of the tree, having cut off all the leaves, and cuts it off to make the Hydra incapable of regeneration—except by growing as a similar tree again (Lucio and Pagli, *op. cit.*, p. 133), and letting the whole process restart, again (which can, *in theory*, go on forever—at this point the labours of Hercules become a one-person game or a puzzle, Turing 1954);

[5] I equate, formally, the *termination* of an algorithm with the *solvability* of the *halting problem for a Turing Machine* (see below).

[6] I am in complete agreement with Jervell (1985, p. 431) (italics added), when he writes: 'It is of course problematic about *what is a natural mathematical theorem*'.

[7] This list can be written semi-formally as Rabin (1957) has done. In this case, the condition for Hercules to satisfy, for an arithmetically *coded* battle (or to complete a task), so that the Hydra is guaranteed to be annihilated, is that 'he' [Hercules] must enumerate a *Post simple set*.

- Thus, depending on the mechanism of regeneration, the game is a combination of a concatenation of an *arithmetical game* and *a puzzle*, the concatenation repeated either finitely or indefinitely[8];
- Both in the *arithmetical game* and in its combination with the *task*, Hercules is the victor—i.e., completes, successfully, the second of his labours or his task;
- However, it must be remembered, at least in the Hydra-Hercules battle, the latter never mis-cuts (the leaf he intends to cut) or chops-off the wrong-head (or no-head at all—which means also no-leaf at all is cut), or due to the obvious *human* trait of *tiredness* allows the untiring, non-human, Hydra[9] to prevail;

Under these conditions, a result of Rabin (*op.cit.*), *effectivising* the Gale-Stewart game (1953) applies to this battle between Hydra and Hercules[10] (theorem 1, p. 152); in *intuitive* terms (p. 148; italics added):

> [T]there are games in which the player who in *theory* can always win, cannot do so in *practice* because it is impossible to supply him with *effective* instructions regarding how he should play in order to win.

In the game between the Hydra and Hercules (and when it becomes a puzzle), Hercules always wins (or completes his task), but no *effective* policy can be recommended by anyone advising him; this is the same as the failure of *programming a computer*—i.e., writing an *algorithm*—to follow the rules specified by it, to win. In other words, the *halting problem* for *any* Turing Machine is *recursively unsolvable* (Davis 1958, Theorem 2.2, p. 70)—i.e., *no* termination program can be devised for this two-person (or one-person) game.

Remark 1 *Effectivity* is defined by equating an *informal* with a *formal* concept; therefore, *computability theory* of the Post-Turing type (Soare 2016) has affinities with the Brouwer-Bishop notion of *constructive mathematics*.[11] However, unrestrained classical mathematics is used in the former, to derive the equality between the informal and the formal (Davis et al. 1976, p. 340).

Remark 2 Actual policy makers do not equate the informal with the formal; this may mean that they succeed in implementing policies that have no algorithmic equivalent—a priori. This is akin to the *negative solution to Hilbert's Tenth Problem*—which does not mean that some Diophantine equations cannot be solved algorithmically, but that there does not exist a general method for their solution.

Remark 3 There is a plethora of results on the existence of termination programs—or, the solving of the *Halting Problem for Turing Machines*—using Set Theory, Model Theory, Proof Theory and Recursion Theory, in terms of finding a way to show the equivalence between the formal and the informal in mathematical logic

---

[8] Neither the Hydra, nor Hercules, seem to get 'tired' or 'bored' with the repeated concatenations.

[9] The Hydra is a non-human monster (see, *eg*., the title and contents of Luccio & Pagli, *op.cit.*).

[10] Even after the arithmetic game becomes a one-person game (Smullyan, *op.cit.,* p. 86) – i.e., quasi-Sisyphean *task*, for which the Smullyan result applies and, therefore, also the Rabin theorem.

[11] Hence, the Goodstein quote in the title page of this paper (but see Sieg 1997).

and the foundations of mathematics. In a sense, the existence of a winning strategy must go *beyond the finite*, in all of the four fields—endorsed by Goodstein.[12]

2.3 The Hydra-Hercules Battle as a *Busy Beaver Game*

> The examples of *non-computable functions* … will be well defined in an extremely primitive sense; we shall use *only* the principle that a non-empty finite set of non-negative integers has a largest element. …. [W]e shall *not* use an enumeration of computable functions to show that our examples are non-computable functions. Thus, in this sense, we do *not* use the *diagonal process*.
>
> **Rado**, *op. cit.,* p. 877; italics added.

As a *homage* to the title of this *Festschrift* in honour of an outstanding scholar, I want to reinterpret the Hydra-Hercules battle as a Busy Beaver Game and, therefore, this section is based on Rado (*op.cit*.) and House and Rado 1964—the latter is *An Approach to Artificial Intelligence* using human intuition, imagination and experience, in terms of the intractability of the uncomputable Busy Beaver function.

I make three changes to the traditional assumptions:

1. In accordance with the eightfold[13] list characterizing the game between Hydra and Hercules, I allow the last in the list to be a possibility; i.e., the game can go on forever or, even, allow for the possibility of a Hydra-win!
2. The change in one of the Rado assumptions for the Busy Beaver Game played with Turing Machine, represented as a sequence of cards, is that the **1**'s listed on the output tape are *unbroken.* In Boolos and Jeffrey 1989, Problem 3.7, p. 27, it is referred to as the *productivity* $p(n)$ of an n-state[14] Turing Machine, which is appropriate in the context of a battle/game between Hydra and Hercules.. Although this is a 'stronger' assumption than Rado makes (in either of the two papers), the implications, particularly in terms of the existence of *uncomputable* functions, are the same.
3. The other change from Rado is that the *halting condition*,[15] if it is ever reached, allows the Hydra to prevail, *as well* (see, 1, above).

[12] See the initial quote, in the title page of this paper.

[13] This reminds me of the noble eight-fold path of Buddhism. I should emphasise that a Hydra-win can only be *pro tempore*!

[14] In the context of Rado's Busy Beaver formalization of a Turing Machine, the states refer to the *number of cards* that defines it.

[15] Indicated by a '**0**' in the final column of the card representation of a Turing Machine.

Two further observation may be made, regarding my Busy Beaver interpretation of the Hydra-Hercules game. The first refers to Smoryński[16] (1980, p. 153; italics added):

> 'If we drop the requirement of *computability*—well, there is Tibor Rado's *Busy Beaver Function*. This irrepressible little fellow grows more rapidly than any *theoretically computable function.*'

To understand the second, within the basic framework of the Busy Beaver game, it may be appropriate to pay attention to the following statement by Rado (*op.cit.,* pp. 877; italics added—see also p. 884):

> [W]e shall use only the principle that *a non-empty finite set of non-negative integers has a largest element*.

In this sense, Ye's *strict finitism* (Ye 2011, especially pp. 51–52), is a congenial aesthetic framework to study the battle/game between Hydra and Hercules; but I am not so sure about equating the activities of the mind with *only* neural processes (pp. 6–10; but see also §1, above). Therefore, I prefer the approach of Rado.

I can, therefore, state a proposition and its corollary:

Proposition 1 No formal procedure can be found for the termination process to be algorithmized.

Corollary 1 There is no algorithmic way to define a victor in the game between the Hydra and Hercules.

Remark 4 Proposition 1 is the Busy Beaver game equivalent of Rabin's Theorem 1 (above).

Remark 5 The 'Proofs' require a discussion and a delineation of the boundaries between imagination, intuition and formalization and the role of axioms in proofs. This is the subject matter of the next section.

2.4 *Notes* on Axioms, Rules of Inference and Proofs in Mathematics and Logic

> A *tree proof* shows the logical structure … more clearly. The topmost node of the tree are *axioms* and at each application of *a rule of inference* we have *a branch.*

[16] Smoryński, in this article, refers to Alan Turing as 'an enigmatic cryptanalyst' (p. 154) and Ramsey as an 'economist' (p. 151)—which is puzzling; I suppose Smoryński does *not* mean that these two were *only* 'cryptanalysts' and 'economists,' they were much more, even by education. In addition, Smoryński claims that the Busy Beaver result on the existence of non-computable functions is due to the diagonalization procedure (pp. 153–4), despite Rado's explicit denial of it (op.cit., p. 877).

> A formula is *provable* if it is the end-formula of a tree proof. The set of provable formulas is recursively enumerable [*re* but not necessarily recursive].
>
> Feferman 1998, p. 180; first and last set of italics in the original.

Proofs in mathematics and (mathematical) logic are, traditionally based on *axioms* and *rules of inference*. *Provability* being the end-formula of a tree-proof is equivalent to a termination of the inference rules (or the solvability of the halting problem for Turing Machines, if *effective algorithms* are *sine quibus non* of provable formulas i.e., *computation rules*, which is *not* traditional). Therefore, *modern proofs* are derived formulas from axioms, rules of inference and computation rules. In this sense, the inadequacy of the Peano Axioms for the proof of the Goodstein sequence or the win-sequence of the Hydra-Hercules game and the need to accept Cantor's theory of infinite ordinals becomes *almost* obvious.

Before I explicate this issue (i.e., Cantor's theory of the infinite ordinals), I must make it clear why I did not use the standard tree-representation of the Hydra-Hercules battle/game. By effectivizing the game, as in Sect. 3, and interpreting it as the Busy Beaver Game in Sect. 4, I have *eschewed reliance on proofs in axioms*—in short, on the computable non-proof of termination rules (or the validity of the non-halting problem for Turing Machines). In my view, proof is an *activity*—a real-life process—for, if not, it cannot be reconciled with the experiments of the life sciences (or any of the sciences, including the social sciences, like economics). Therefore the 'real-life' activity of a Hercules, in chopping-off a head, whatever the regenerative rule adopted by Hydra, and trying to reach a branch with only links to a root, is—for me—the proof that a termination has been achieved, it the provability of a formula, as an activity.

This is (partly) because I adhere to the conjunction of the Brouwer-Bishop notion of constructive mathematics and the Turing-Post concept of computability (the latter define formal systems as *recursively enumerable* sets that are *not* necessarily r*ecursive*). Neither pair need axioms for proofs and provability, because they—the axioms—are intrinsically ambiguous (as the quote from Bishop signifies—where definitions are axioms):

> [I]t is *impossible* to consider *every* possible interpretation of our definition and say whether that is what we have *in mind*. There is *always ambiguity*. The expositor himself can *never fully know* all the possible ramifications of his definitions, and he is subject to the same necessity of modifying his interpretations, and sometimes his definitions as well, to *conform to the dictates of experience*.
>
> Bishop 1967, p. 7; italics added.

This observation by Bishop, as *a constructivist mathematician* (reflecting also Brouwer's standpoint), is exactly (!) what is meant by also in Goodstein's:

> [*N*]*othing can be expressed in a form which is entirely exempt from misunderstanding.*
>
> Goodstein 1962, p. 148; italics in the original.

The former can be considered the *intrinsic*, the latter the *extrinsic*, aspect of a characterisation of *axioms*. Axioms are supposed to summarise properties of concepts[17]—things, entities; however, they can *never* express *all* of the intentions, imaginations and intuitions of a *mind*. As the constructivist mathematician Bishop says, 'there is *always ambiguity*'—and this leads to '*misunderstanding*,' in the words of Goodstein, who was also a constructivist mathematician (but different from Bishop's type). Brouwer, even though he supported Heyting's axiomatic formalization of intuitionistic logic, was fundamentally against *any* kind of axiomatization (in mathematics). His partial concession to logic and language is in Brouwer (1981)—but he never compromised on the autonomy of mathematics and the rules of mental constructions that formed the basis on intuitionistic logic.[18]

Note, however, that these (and many other) constructivists[19] do not differentiate the 'ambiguities' or 'misunderstandings' in terms of belonging to finite or (countably) infinite structures.

I have *not* used the *axiom of determinateness* (explicitly) in making the game interpretation of the Hydra-Hercules battle determined (or decidable), even in a non-algorithmic or non-constructive (particularly not the kind that uses something similar to the Church-Turing thesis, for e.g., Russian/Markov constructivism) sense. Nor have I felt any need, because of the inadequacy of the *Peano Axioms* (of arithmetic), to adhere to Cantor's theory of the infinite and accept ordinals, in addition to cardinals. Even the finite and infinite Ramsey theorems[20] and the Ramsey numbers, ubiquitous in the standard interpretation and analysis of the independence of Peano axioms for the proof of termination of a computable sequence has been eschewed. Instead, I have emphasized effectivity and non-computability in the game between the Hydra and Hercules.

To return to Cantor and infinity, in the context of axioms, I am fundamentally opposed to any form of the actual infinite, as against the potential infinite; if not,

[17] In the traditional Euclidean sense—even if not the 'modern' Hilbertian way—axioms are supposed to be about 'self-evident propositions'—but 'self-evident' *to* whom?

[18] Boole's approach to the mathematization of logic is through a similar aspect of the rules of mental constructions and therefore, I consider him a distinguished predecessor of intuitionistic logic (see, in particular, Brouwer's contribution to Boole's centennial anniversary, Brouwer (1954–56).

[19] I claim that similar views were held by Turing and Feferman. For my purposes it is sufficient to note Feferman's important observation (1998, p. 178; italics added):

'[T]he behavior of logical systems *cannot* be deduced from physical principles and the *creative* and *intuitive* aspects of mathematical work *evade* logical encapsulation'.

It is obvious that I am greatly indebted to Van Atten (2004), particularly Chap. 2, for my views on proofs and logic.

[20] Ramsey himself used the infinite version of his theorem to 'prove' the finite version, using the axiom of choice (he called it the axiom of selection, Ramsey, *op.cit*., p. 82; incidentally, Ramsey was motivated to consider issues in combinatorics, leading to Ramsey theory, by the *Entscheidungsproblem*—which his younger King's College Fellow, Alan Turing, solved a few years later, but only by making the *human element* in the formalization of *effective* 'as clear as mud'—*pace* Meredith 1931).

I cannot ‘walk on two legs.’ Be that as it may, Cantor’s almost vehement[21] rejection of the infinitesimal, Veronese—and his concept of the infinite, based on the rejection of the Archimedean axiom (see Hobson 1927, 1957, p. 58[22]), is based on Hilbert’s axiom of continuity (Dauben 1980, pp. 216–219). This, at least for me, shows clearly the absurdity of rejecting (or accepting) anything on the basis of axioms (the infinitesimals were, happily, resurrected by Abraham Robinson and Edward Nelson).

Moreover, no consideration of *Hilbert’s Hotel* helps in the development of *intuition* when considering actual infinities (Devlin 2006). The clerk, in Devlin’s Angle, attended a course in mathematics of the absurd—for no one can ‘ask every guest to move to the room with the room number the next integer’ for a countable (actual) infinity in ‘a minute [of making] phone calls.’ A machine or a human being can only develop intuition on the basis of potential infinities.

That proofs, particularly mathematical proofs, are important is, of course, folklore in intuitionistic (even non-intuitionistic) constructive mathematics—as even a cursory glance of Brouwer, Goodstein and Bishop will vouchsafe. Their kind of mathematical proof can be—is—devoid (or can be made independent) of axioms. This is the kind of proof that House and Rado (*op.cit*., pp. 115–116) consider relevant for artificial intelligence via any Busy Beaver game. *A fortiori*, non-termination rules, regeneration rules, etc., in the typical Hydra-Hercules game, may or may not terminate.

2.5 Concluding Notes

> *Artificial Intelligence* has a habit (already demonstrated for psychology and linguistics, among other areas) of overturning or at least disturbing the *standard outlooks* in subjects with which it comes in contact. There is no reason to suppose that *mathematical logic* will be immune to this process.
>
> **Campbell**, 1987, p. 529; italics added.

I consider, on the one hand, ‘walking on two feet’ in terms of the active, real-life, proof of termination in the Hydra-Hercules game; the auxiliary consideration of many virus’s eternal lives, as analogous to the Hydra’s regeneration rules, is not part of the main story. Hercules winning—even using the best possible strategy—is only one of the possible alternative outcomes. Therefore, this view of proof as an applied, active, part of life—given to intuitive aspects of humans as well as machines, completely independent of axioms—is ‘one foot’ in the life of *artificially intelligent human beings*.

---

[21] I nearly said ‘vicious’—instead of vehement!

[22] Hobson refers to the axiom of Archimedes as the ‘principle of Archimedes,’ in *op.cit*., p. 58.

Fig. 2.2 Shu (right) with Ragupathy (left), Connie Wang (middle) and Selda Kao[23] next to Vela, Taipei, September, 2012

However, the Hydra-Hercules game, first as an uncomputable arithmetic game and, then, in terms of the uncomputable Busy Beaver, forms the theoretical foot of the exercise. It is the artificial intelligence of machine life.

Therefore, the two feet on which I walk, along the path broached by the wise Chinese, are both based on artificial intelligence, but one human-thinking and the other machine-learning. In this way, *I fuse humans and machines at the same time as thinking and learning, by artificial intelligence.*

It was essential, for this fusing to 'stick' that I divest proof theory of all vestiges of axioms and make it fully dependent on inference and computation rules; the novelty here is, on the one hand, exploiting the human element—for example *intuition*—in effective processes. The other, equally important factor is ascribing intuition and talking about effective inference rules. For too long, inference rules have been assumed to be 'objective' and independent of the human element (Fig. 2.2).

A 'model' for my two-pronged approach—walking on two feet—is the Appel-Haken proof of the celebrated four-colour theorem. The proof that any planar-map is four-colourable requires the collaboration of 'blackboard- and -chalk' and computers; i.e., humans and machines—with ample scope for intuitions and learning, via the mediating force of a kind of artificial intelligence.

It is with the utmost humility that this essay is penned, in honour of a scholar of impeccable credentials and humane attributes. I cannot help thinking of *Samuel Beckett*, in **Worstward Ho**, saying (in 1983):

[23] Selda Kao is Mrs. Ragupathy and a former student of Shu and Vela; Connie Wang is one of the forces behind this *Festschrift* in honour of Shu-Heng Chen.

'Ever tried. Ever failed. No matter. Try again. Fail again. Fail better.' I have tried; if I have failed, I hope I failed better!

References

Ashby WR (1945) Effect of controls on stability. Nature 155:242–243

Bishop E (1967) Foundations of constructive analysis. McGraw-Hill Inc., New York

Boolos GS, Jeffrey RC (1989) Computability and logic, 3rd edn. Cambridge University Press, Cambridge

Brouwer LEJ (1954–56) The effect of intuitionism on classical algebra of logic. In: Proceedings of the Royal Irish Academy. Section a: mathematical and physical sciences, vol 57, pp 113–116

Brouwer LEJ (1981) Brouwer's Cambridge lectures on Intuitionism, van Dalen D (ed). Cambridge University Press, Cambridge

Cajal SRY (1937, 1989) Recollections of my life. The MIT Press, Cambridge

Chomsky N (1957) Syntactic structures. Mouton & Co, Ltd, The Hague

Dauben JW (1980) The development of cantorian set theory, chapter 5. In: Grattan-Guinness I (ed) From the calculus to set theory**,** 1630 – 1910: an introductory history. Princeton University Press, Princeton, pp 181–219

Davis M (1958) Computability and unsolvability. McGraw-Hill Book Company, New York

Davis M, Matiyasevič Y, Robinson J (1976) Hilbert's tenth problem. diophantine equations: positive aspects of a negative solution. In: Proceedings of symposia in Pure mathematics, mathematical developments arising from hilbert problems, vol XXVIII, Part 2. American Mathematical Society, Providence, pp 372–378

Devlin K (2006) Devlin's angle: infinity and intuition, Jan, downloaded from the Internet on 1 Sep 2020

Feferman S (1998) In the light of logic. Oxford University Press, Oxford

Gale D, Stewart FM (1953). Infinite games with perfect information, chap 13. In: Kuhn HW, Tucker AW (eds) Contributions to the theory of games, vol II, Annals of mathematics studies, vol 28. Princeton University Press, Princeton, pp 245–266

Gardner M (1983) Mathematical Games. Sci Am 249(2):12–21

Gödel K (1931, 1986). On formally undecidable propositions of Principia Mathematica and related systems I. In: Solomon F et al (ed) Kurt Gödel—collected works, vol 1. Publications 1929–1936. Oxford University Press, Oxford, pp 144–195

Gold EM (1964) Language identification in the limit, research memoranda, RM-4136-PR. RAND Corporation, Santa Monica, Los Angeles

Goodstein RL (1962) The axiomatic method. Proc Aristot Soc Suppl 36:145–154

Goodstein RL (1944) On the restricted ordinal theorem. J Symb Log 9(2):33–41

Hebb DO (1949) The organization of behavior: a neurophysiological theory. Wiley Inc., New York

Hobson EW (1927, 1957) The theory of functions of a real variable and the theory of fourier series, vol 1. Dover Publications Inc, New York

Hobson EW (1927) The theory of functions of a real variable & the theory of Fourier's series. Cambridge University Press, Cambridge

Hodgson B (2004) Herculean or Sisyphean tasks? European Mathematical Society (EMS), Newsletter, vol 51, pp 11–16

House RW, Rado T (1964) An approach to artificial intelligence. IEEE Trans Commun Electron (70):111–116

Husbands P, Holland O (2008) The ratio club: a hub of british cybernetics, chap 6. In: Holland O, Husbands P , Wheeler M (eds) The mechanical mind in history. The MIT Press, Cambridge, pp. 91–148

Jervell HR (1985) Gentzen games. Math Logic Q 31(25–28):431–439

Kirby L, Paris J (1982) Accessible independence results for Peano arithmetic. Bull Lond Math Soc 14(4):285–293

Luccio F, Pagli L (2000) Death of a monster. ACM SIGACT News 31(4):130–133

McCall JJ (1989) The Smithian self and It's Bayesian brain, WP 596, Department of Economics, UCLA

McCulloch WS, Pitts W (1943) A Logical calculus of the ideas immanent in nervous activity. Bullet Math Biophy 5:115–133

Meredith HO (1931) Letter to Keynes. In: Donald EM (ed) The collected writings of John Maynard Keynes. The general theory and after-part i: preparation, vol XIII. The Macmillan Press Ltd, London, pp 266–267

Paris J, Harrington L (1977) A mathematical incompleteness in Peano arithmetic, chapter D8. In: Barwise J (ed) Handbook of mathematical logic. North-Holland Publishing Company, Amsterdam, pp 1133–1142

Rabin MO (1957) Effective computability of winning strategies. In: Dresher M, Tucker AW, Wolfe P (eds) Annals of mathematics studies, No. 39, Contributions to the theory of games, vol III. Princeton University Press, Princeton

Rado T (1962) On non-computable functions. Bell Syst Tech J 41(3):877–884

Ramsey FP (1931) The foundations of mathematics and other logical essays. In: Braithwaite RB. Routledge, London

Rosenbleuth A, Norbert W, Bigelow J (1943) Behavior, purpose, and teleology. Philos Sci 10(1):18–24

Rumelhart DE, James LM, The PDP Research Group (1986a) Parallel distributed processing: explorations in the microstructure of cognition —foundations, vol 1. The MIT Press, Cambridge

Rumelhart DE, James LM, The PDP Research Group (1986b) Parallel distributed processing: explorations in the microstructure of cognition—psychological and biological models, vol 2. The MIT Press, Cambridge

Sherrington CS (1906) The integrative action of the nervous system. Charles Scribner's Sons, New York

Sherrington SC (1940) Man on his Nature: The Gifford lectures, Edinburgh, 1937–38. Cambridge University Press, Cambridge

Sieg W (1997) Step by recursive step: church's analysis of effective calculability. Bull Symb Log 3(2):154–180

Smullyan RM (1979) Trees and ball games, annals of the New York academy of sciences, vol 321, issue 1, pp 86–90

Smoryński C (1980) Some raidly growing functions, mathematical intelligencer, vol 2, September 1980, pp 149–154

Soare RI (2016) Turing computability: Theory and applications. Springer, Heidelberg

Turing A (1936–37) On computable numbers, with an application to the Entscheidungsproblem. In: Proceedings of the London Mathematical Society, Series 2, vol 42, issue 1, pp 230–265

Turing AM (1954) Solvable and unsolvable problems. Sci News 31:7–23

Turing AM (1948, 1992) Intelligent Machinery. In: Ince DC (ed) Collected works of A.M. Turing: mechanical intelligence. North-Holland, Amsterdam, pp 107–127

Van Atten M (2004) On brouwer. Thomson/Wadsworth, Thomson learning, London

Wiener N (1948, 1961) Cybernetics–or the control and communication in the animal and the machine

Ye F (2011) Strict finitism and the logic of mathematical applications. Synthese library, vol 355. Springer, London

Chapter 3
Logic and Epistemology in Behavioral Economics

J. Barkley Rosser Jr.

Abstract Shu-Heng Chen was a deep student agent-based computational economics. He considered underlying issues arising with this approach including how it relates to experimental economics and behavioral economics. Herbert Simon founded behavioral economics and developed the concept of bounded rationality arising from the limits of knowledge of agents, the epistemological problem. He saw these as arising both from lack of knowledge but also from deeper computational problems arising from logical paradoxes associated with incompleteness theorems that imply non-computability of systems. Problems of knowledge also arise in such systems when they involve nonlinear dynamics leading to various forms of dynamic complexity involving chaos theory and fractal dimensionalities. This paper considers these issues in connection with the work of Chen.

Keywords Agent-based modeling · Logic · Epistemology · Behavioral economics · Complex dynamics

3.1 Introduction

Shu-Heng Chen has been a deep student of agent-based computational economics (2016). This has involved various applications such as the behavior of cobweb dynamics (Chen and Yeh 1996), financial markets (Chen and Yeh 1997, 2002), macroeconomics (Chen 2003), and the design of lottery markets (Chen and Chie 2008), among others. These efforts have led him to consider deeper implications of for the nature of economic agents (Chen 2012) and how their decisionmaking happens in their brains (Chen 2014). This has led him to consider the relation between computational economics, experimental economics, and behavioral economics, with a perspective that draws substantially on ideas of the founder of behavioral economics, Herbert Simon (Chen 2005; Chen and Kao 2016), with Chen having at times declared his work to be "in the tradition" of Simon.

J. B. Rosser Jr. (✉)
JMU College of Business, James Madison University, Harrisonburg, VA 22807, USA
e-mail: rosserjb@jmu.edu

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_3

Knowledge is hard to obtain regarding complicated reality and complicated systems. However, complex systems lead to even greater problems of knowledge than complicated ones, even when in important ways complex systems may appear simpler than complicated ones. A complicated system will have many parts that are interconnected in a variety of ways that may not be obvious and may be hard to discern or untangle. However merely complicated systems will "add up" in a reasonably straightforward way. Once one figures out these interconnections and their nature, one can understand the whole relatively easily as it will ultimately be the sum of those parts, which may nevertheless be hard to understand on their own. However, in the case of complex systems, by their nature they usually manifest that phenomenon first identified by Aristotle that the whole may be greater than the sum of the parts. This greater degree of wholeness will often be due to nonlinear relations within the system such as increasing returns to scale or tangled non-monotonic relations. Even though there may be fewer variables and relations, the complex nature of the relations makes knowledge and understanding of the system more difficult (Israel 2005).

The knowledge problem is more formally known as the epistemological problem, and this author has previously addressed it in this context (Rosser 2004). However, while drawing on discussion in that paper, this one will not only update arguments made there but will consider additional topics such as the complexity foundations of Herbert Simon's "bounded rationality" concept (Simon 1957, 1962).

How nonlinear dynamical systems manifest problems of knowledge is easily seen for chaotic systems, which are characterized by the problem of sensitive dependence on initial conditions, known popularly as the "butterfly effect." If minute changes in initial conditions, either of parameter values controlling a system or of initial starting values, can lead to very large changes in subsequent outcomes of a system, then it may essentially require an infinite precision of knowledge to completely know the system, which undermines the possibility for rational expectations for such systems (Rosser 1996). Also, fractality of dynamic attractor basin boundaries in systems with multiple such basins can behave similarly such that even the slightest amount of stochastic noise in the dynamical system can lead to very different outcomes (Rosser 1991).

The problem of logic or computation arises in complex systems of multiple interacting heterogeneous agents thinking about each others' thinking. Although game theoretic solutions such as Nash equilibria may present themselves, these may involve a certain element of ignorance, a refusal to fully know the system. Efforts to fully know the system may prove to be impossible due to problems of infinite regress or self referencing that lead to non-computability (Binmore 1987; Albin with Foley 1998; Koppl and Rosser 2002; Mirowski 2002; Landini et al. 2020; Rosser 2021). This becomes entangled with deeper problems in the foundations of mathematics involving constructivist logic and its link to computability (Velupillai 2000; Zambelli 2004; Rosser 2010, 2012, 2021; Kao et al. 2012).

We consider the role of Herbert Simon in understanding the deep relations between complexity and the limits of knowledge. As a founding figure in the study of artificial intelligence, he was fully aware of the computational complexity issues arising from the logical paradoxes of self-referencing and related matters. He was also of the limits of computational capabilities of humans as well as the high cost of obtaining information. From these ideas he developed the concept of *bounded rationality* (Simon 1957) as he basically invented modern behavioral economics based on this. He also dug more deeply into complexity issues as he also largely developed the idea of hierarchical complexity (Simon 1962), which adds further layers to the epistemological difficulties associated with understanding complex systems. The influence of these ideas of Simon has been both deep and wide (Rosser and Rosser 2015).

3.2 Forms of Complexity?[1]

In *The End of Science* (1996, p. 303) John Horgan reports on 45 definitions of complexity that have been gathered by the physicist Seth Lloyd. Some of the more widely used conceptualizations include *informational entropy* (Shannon 1948), *algorithmic complexity* (Chaitin 1987), *stochastic complexity* (Rissanen 1989), and *hierarchical complexity* (Simon 1962). Three other definitions have been more frequently used in economics that do not appear on Lloyd's list, namely simple complicatedness in the sense of many different sectors with many different interconnections (Pryor 1995; Stodder 1995), dynamic complexity (Day 1994), and computational complexity (Lewis 1985; Velupillai 2000). We shall not consider the knowledge problems associated with mere complicatedness, which are simpler than those associated with true complexity (Israel 2005).

According to Day (1994), dynamic complexity can be defined as arising from dynamical systems that endogenously fail to converge to either a point, a limit cycle, or a smooth explosion or implosion. Nonlinearity is a necessary but not sufficient condition for such complexity. Rosser (1999) identifies this definition with a *big tent* view of dynamic complexity that can be subdivided into four sub-categories: *cybernetics, catastrophe theory, chaos theory,* and *small tent complexity* (now more commonly called *agent-based complexity*) The latter does not possess a definite definition, however Arthur et al. (1997) argue that such complexity exhibits six characteristics: (1) dispersed interaction among locally interacting heterogeneous agents in some space, (2) no global controller that can exploit opportunities arising from these dispersed interactions, (3) cross-cutting hierarchical organization with many tangled interactions, (4) continual learning and adaptation by agents, (5) perpetual novelty in the system as mutations lead it to evolve new ecological niches, and (6) out-of-equilibrium dynamics with either no or many equilibria and little likelihood of a global optimum state emerging. Certainly such systems offer considerable scope for problems of how to know what is going on in them.

Computational complexity essentially amounts to a system being non-computable. Ultimately this depends on a logical foundation, that of non-recursiveness due to incompleteness in the Gödel sense (Church 1936; Turing 1937).

[1] Substantial portions of this discussion draw on Rosser (2021).

In actual computer problems this problem manifests itself most clearly in the form of the halting problem (Blum et al. 1998), that the halting time of a program is infinite. Ultimately this form of complexity has deep links with several of the others listed above, such as Chaitin's algorithmic complexity. These latter two approaches are the ones we shall consider in more detail in the next two sections.

3.3 Dynamic Complexity and Knowledge

In dynamically complex systems, the knowledge problem becomes the general epistemological problem. Consider the specific problem of being able to know the consequences of an action taken in such a system. Let $G(\mathbf{x}_t)$ be the dynamical system in an n-dimensional space. Let an agent possess an action set $\mathbf{A}$. Let a given action by the agent at a particular time be given by $\mathbf{a}_{it}$. For the moment let us not specify any actions by any other agents, each of whom also possesses his or her own action set. We can identify a relation whereby $\mathbf{x}_t = f(\mathbf{a}_{it})$. The knowledge problem for the agent in question thus becomes, "Can the agent know the reduced system $G(f(\mathbf{a}_{it})$ when this system possesses complex dynamics due to nonlinearity"?

First of all, it may be possible for the agent to be able to understand the system and to know that he or she understands it, at least to some extent. One reason why this can happen is that many complex nonlinear dynamical systems do not always behave in erratic or discontinuous ways. Many fundamentally chaotic systems exhibit *transiency* (Lorenz 1992). A system can move in and out of behaving chaotically, with long periods passing during which the system will effectively behave in a non-complex manner, either tracking a simple equilibrium or following an easily predictable limit cycle. While the system remains in this pattern, actions by the agent may have easily predicted outcomes, and the agent may even be able to become confident regarding his or her ability to manipulate the system systematically. However, this essentially avoids the question.

Let us consider four forms of complexity: chaotic dynamics, fractal basin boundaries, discontinuous phase transitions in heterogeneous agent situations, and catastrophe theoretic models related to this third form. For the first of these there is a clear problem for the agent, the existence of sensitive dependence on initial conditions. If an agent moves from action $\mathbf{a}_{it}$ to action $\mathbf{a}_{jt}$, where $|\mathbf{a}_{it} - \mathbf{a}_{jt}| < \varepsilon < 1$, then no matter how small ε is, there exists an m such that $|G(f(\mathbf{a}_{it+t'}) - G(f(\mathbf{a}_{jt+t'})| > m$ for some t' for each ε. As ε approaches zero, m/ε will approach infinity. It will be very hard for the agent to be confident in predicting the outcome of changing his or her action. This is the problem of the butterfly effect or sensitive dependence on initial conditions. More particularly, if the agent has an imperfectly precise awareness of his or her actions, with the zone of fuzziness exceeding ε, the agent faces a potentially large range of uncertainty regarding the outcome of his or her actions. In Edward Lorenz's (1963) original study of this matter when he "discovered chaos," when he restarted his simulation of a three-equation system of fluid dynamics partway through, the roundoff error that triggered a subsequent dramatic divergence was too small for his computer to "perceive" (at the four decimal place).

There are two offsetting elements for chaotic dynamics. Although an exact knowledge is effectively impossible, requiring essentially infinitely precise knowledge (and knowledge of that knowledge), a broader approximate knowledge over time may be possible. Thus, chaotic systems are generally bounded and often ergodic (although not always). While short-run relative trajectories for two slightly different actions may sharply diverge, the trajectories will at some later time return toward each other, becoming arbitrarily close before once again diverging. Not only may the bounds of the system be knowable, but the long-run average of the system may be knowable. There are still limits as one can never be sure that one is not dealing with a long transient of the system, with it possibly moving into a substantially different mode of behavior later. But the possibility of a substantial degree of knowledge, with even some degree of confidence regarding that knowledge is not out of the question for chaotically dynamic systems.

Regarding fractal basin boundaries, first identified for economic models by Hans-Walter Lorenz (1992) in the same paper in which he discussed the problem of chaotic transience. Whereas in a chaotic system there may be only one basin of attraction, albeit with the attractor being fractal and strange and thus generating erratic fluctuations, the fractal basin boundary case involves multiple basins of attraction, whose boundaries with each other take fractal shapes. The attractor for each basin may well be as simple as being a single point. However, the boundaries between the basins may lie arbitrarily close to each other in certain zones.

In such a case, although it may be difficult to be certain, for the purely deterministic case once one is able to determine which basin of attraction one is in, a substantial degree of predictability may ensue, although again there may be the problem of transient dynamics, with the system taking a long and circuitous route before it begins to get anywhere close to the attractor, even if the attractor is merely a point in the end. The problem arises if the system is not strictly deterministic, if G includes a stochastic element, however small. In this case one may be easily pushed across a basin boundary, especially if one is in a zone where the boundaries lie very close to one another. Thus there may be a sudden and very difficult to predict discontinuous change in the dynamic path as the system begins to move toward a very different attractor in a different basin. The effect is very similar to that of sensitive dependence on initial conditions in epistemological terms, even if the two cases are mathematically quite distinct.

Nevertheless, in this case as well there may be something similar to the kind of dispensation over the longer run we noted for the case of chaotic dynamics. Even if exact prediction in the chaotic case is all but impossible, it may be possible to discern broader patterns, bounds and averages. Likewise in the case of fractal basin boundaries with a stochastic element, over time one should observe a jumping from one basin to another. Somewhat like the pattern of long run evolutionary game dynamics studied by Binmore and Samuelson (1999), one can imagine an observer keeping track of how long the system remains in each basin and eventually developing a probability profile of the pattern, with the percent of time the system spends in each basin possibly approaching asymptotic values. However, this is contingent on the nature of the stochastic process as well as the degree of complexity of the fractal pattern of the basin boundaries. A non-ergodic stochastic process may render it very difficult, even impossible, to observe convergence on a stable set of probabilities for being in the respective basins, even if those are themselves few in number with simple attractors.

For the case of phase transitions in systems of heterogeneous locally interacting agents, the world of the so-called "small tent complexity." Brock and Hommes (1997) have developed a useful model for understanding such phase transitions, based on statistical mechanics. This is a stochastic system and is driven fundamentally by two key parameters, a strength of interactions or relationships between neighboring agents and a degree of willingness to switch behavioral patterns by the agents. For their model the product of these two parameters is crucial, with a bifurcation occurring for their product. If the product is below a certain critical value, then there will be a single equilibrium state. However, once this product exceeds a particular critical value two distinct equilibria will emerge. Effectively the agents will jump back and forth between these equilibria in herding patterns. For financial market models (Brock and Hommes 1998) this can resemble oscillations between optimistic bull markets and pessimistic bear markets, whereas below the critical value the market will have much less volatility as it tracks something that may be a rational expectations equilibrium.

For this kind of a setup there are essentially two serious problems. One is determining the value of the critical threshold. The other is understanding how the agents jump from one equilibrium to the other in the multiple equilibrium zone. Certainly the second problem resembles somewhat the discussion from the previous case, if not involving as dramatic a set of possible discontinuous shifts.

Of course once a threshold of discontinuity is passed it may be recognizable when it is approached again. But prior to doing so it may be essentially impossible to determine its location. The problem of determining a discontinuity threshold is a much broader one that vexes policymakers in many situations, such as attempting to avoid catastrophic thresholds that can bring about the collapse of a species population or of an entire ecosystem. One does not want to cross the threshold, but without doing so, one does not know where it is. However, for less dangerous situations involving irreversibilities, it may be possible to determine the location of the threshold as one moves back and forth across it.

On the other hand in such systems it is quite likely that the location of such thresholds may not remain fixed. Often such systems exhibit an evolutionary self-organizing pattern in which the parameters of the system themselves become subject to evolutionary change as the system moves from zone to zone. Such non-ergodicity is consistent not only with Keynesian style uncertainty, but may also come to resemble the complexity identified by Hayek (1948, 1967) in his discussions of self-organization within complex systems. Of course for market economies Hayek evinced an optimism regarding the outcomes of such processes. Even if market participants may not be able to predict outcomes of such processes, the pattern of self-organization will ultimately be largely beneficial if left on its own. Although Keynesians and Hayekian Austrians are often seen as in deep disagreement, some observers have noted the similarities of viewpoint regarding these underpinnings of uncertainty (Shackle 1972; Loasby 1976; Rosser 2021). Furthermore, this approach leads to the idea of the openness of systems that becomes consistent with the critical realist approach to economic epistemology (Lawson 1997).

Considering this problem of important threshold brings us to the final of our forms of dynamic complexity to consider here, catastrophe theory interpretations. The knowledge problem is essentially that previously noted, but is more clearly writ large as the discontinuities involved are more likely to be large as the crashes of major speculative bubbles. The Brock-Hommes model and its descendants can be seen as a form of what is involved, but returning to earlier formulations brings out underlying issues more clearly.

The very first application of catastrophe theory in economics by Zeeman (1974) indeed considered financial market crashes in a simplified two-agent formulation: fundamentalists who stabilized the system by buying low and selling high and "chartists" who chase trends in a destabilizing manner by buying when markets rise and selling when they fall. As in the Brock-Hommes formulation he allows for agents to change their roles in response to market dynamics so that as the market rises fundamentalists become chartists, accelerating the bubble, and when the crash comes they revert to being fundamentalists, accelerating the crash. Rosser (1991) provides an extended formalization of this in catastrophe theory terms that links it to the analysis of Minsky (1972) and Kindleberger (1978), further taken up in Rosser et al. (2012) and Rosser (2020a). This formulation involves a cusp catastrophic formulation with the two control variables being the demands by the two categories of agents, with the chartists' demand determining the position of the cusp that allows for market crashes.

The knowledge problem here involves something not specifically modeled in Brock and Hommes, although they have a version of it. It is the matter of the expectations of agents about the expectations of the other agents. This is effectively the "beauty contest" issue discussed by Keynes in Chap. 12 of this *General Theory* (1936). The winner of the beauty contest in a newspaper competition is not who guesses the prettiest girl, but who guesses best the guesses of the other participants. Keynes famously noted that one could start playing this about guessing the expectations of others in their guesses of others' guesses, and that this could go to higher levels, in principle, an infinite regress leading to an impossible knowledge problem. In contrast, the Brock and Hommes approach simply has agents shifting strategies after watching what others do. These potentially higher level problems do not enter in. These sorts of problems reappear in the problems associated with computational complexity.

3.4 Knowledge Problems of Computational Complexity

Regarding computational complexity, Velupillai (2000) provides definitions and general discussion and Koppl and Rosser (2002) provide a more precise formulation of the problem, drawing on arguments of Kleene (1967), Binmore (1987), Lipman (1991), and Canning (1992). Velupillai defines computational complexity straightforwardly as "intractability" or insolvability. Halting problems such as studied by Blume et al. (1998) provide excellent examples of how such complexity can arise, with this problem first studied for recursive systems by Church (1936) and Turing (1937).

In particular, Koppl and Rosser reexamined the famous "Holmes-Moriarty" problem of game theory, in which two players who behave as Turing machines contemplate a game between each other involving an infinite regress of thinking about what the other one is thinking about. This has a Nash equilibrium, but "hyper-rational" Turing machines cannot arrive at knowing it has that solution or not due to the halting problem. That the best reply functions are not computable arises from the self-referencing problem involved fundamentally similar to those underlying the Gödel Incompleteness Theorem (Rosser , 1936; Kleene 1967, p. 246). Such problems extend to general equilibrium theory as well (Lewis 1991; Richter and Wong 1999; Landini et al. 2020).

Binmore's (1987, pp. 209–212) response to such undecidability in self-referencing systems invokes a "sophisticated" form of Bayesian updating involving a degree of greater ignorance. Koppl and Rosser agree that agents can operate in such an environment by accepting limits on knowledge and operate accordingly, perhaps on the basis of intuition or "Keynesian animal spirits" (Keynes 1936). Hyper-rational agents cannot have complete knowledge, essentially for the same reason that Gödel showed that no logical system can be complete within itself.

However, even for Binmore's proposed solution there are also limits. Thus, Diaconis and Freedman (1986) have shown that Bayes' Theorem fails to hold in an infinite dimensional space. There may be a failure to converge on the correct solution through Bayesian updating, notably when the basis is discontinuous. There can be convergence on a cycle in which agents are jumping back and forth from one probability to another, neither of which is correct. In the simple example of coin tossing, they might be jumping back and forth between assuming priors of 1/3 and 2/3 without ever being able to converge on the correct probability of 1/2. Nyarko (1991) has studied such kinds of cyclical dynamics in learning situations in generalized economic models.

Koppl and Rosser compare this issue to that of the Keynes's problem (1936, Chap. 12) of the beauty contest in which the participants are supposed to win if they most accurately guess the guesses of the other participants, potentially involving an infinite regress problem with the participants trying to guess how the other participants are going to be guessing about their guessing and so forth. This can also be seen as a problem of reflexivity (Rosser 2020b). A solution comes by in effect choosing to be somewhat ignorant or boundedly rational and operating at a particular level of analysis. However, as there is no way to determine rationally the degree of boundedness, which itself involves an infinite regress problem (Lipman 1991), this decision also ultimately involves an arbitrary act, based on animal spirits or whatever, a decision ultimately made without full knowledge.

A curiously related point here is the newer literature (Gode and Sunder 1993; Mirowski 2002) on the behavior of zero intelligence traders. Gode and Sunder have shown that in many artificial market setups zero intelligence traders following very simple rules can converge on market equilibria that may even be efficient. Not only may it be necessary to limit one's knowledge in order to behave in a rational manner, but one may be able to be rational in some sense while being completely without knowledge whatsoever. Mirowski and Nik-Kah (2017) argue that this completes a transformation of the treatment of knowledge in economics in the post-World war II era from assuming that all agents have full knowledge to all agents having zero knowledge.

A further point on this is that there are degrees of computational complexity (Velupillai 2000; Markose 2005), with Kolmogorov (1965) providing a widely accepted definition that the degree of computational complexity is given by the minimum length of a program that will halt on a Turing machine. We have been considering the extreme cases of no halting, but there is indeed an accepted hierarchy among levels of computational complexity, with the knowledge difficulties experiencing qualitative shifts across them. At the lowest level are linear systems, easily solved, with such a low level of computational complexity we can view them as not complex. Above that level are polynomial (P) problems that are substantially more computationally complex, but still generally solvable. Above that are exponential and other non-polynomial (NP) problems that are very difficult to solve, although it remains as yet unproven that these two levels are fundamentally distinct, one of the most important unsolved problems in computer science. Above this level is that of full computational complexity associated where the minimum length is infinite, where the programs do not halt, the sort we have discussed in most of this section. Here the knowledge problems can only be solved by becoming effectively less intelligent.

3.5 Complexity Foundations of Bounded Rationality and Limited Knowledge

Herbert A. Simon was a polymath who published over 900 papers in numerous disciplines and is generally considered to be the "father of modern behavioral economics" (Rosser and Rosser 2015). He certainly coined the term (Simon 1955), although earlier economists certainly accepted many ideas of behavioral economics going at least as far back as Adam Smith (1759). Central to his conception of behavioral economics was the concept of *bounded rationality*. His concern with this idea and his search for its ultimate foundations would lead him to consider the "thinking" of computers as a way of studying human thinking, with this making him a founder of the field of artificial intelligence (Simon 1969).

What is not widely recognized is how complexity ideas underlie this fundamental idea of Simon's. He was fully aware of the debates in logic regarding the solvability of recursive systems (Church 1936; Rosser 1936; Turing 1937) and indeed the deeply underlying problems of incompleteness and inconsistency that hold for any computational system whether one in a fully rational person's head or inside a computer. The limits imposed by computational complexity were for him profound and ultimate. However, even before these limits were reached he doubted the computational capabilities of humans at more basic levels, especially in the face of a reality full of complex systems. And Simon was aware of the unusual probability distributions that nonlinear dynamical systems can generate (Ijiri and Simon 1964). In addition, his awareness of hierarchical complexity simply added to his understanding of the limits of knowledge by the many forms of complexity (Simon 1962), with Simon one of the few figures so early on to be attuned to the multiple varieties of complexity.

Simon's awareness of the limits to knowledge and the importance of bounded rationality led to him emphasizing various important concepts. Thus he distinguished *substantive* from *procedural* rationality (Simon 1969), with the latter what boundedly rational agents due in the face of the limits to their knowledge. They adopt heuristic rules of thumb, and knowing that they will be unable to fully optimize, they seek to achieve set goals, *satisficing*, with Simon's followers developing this into a whole theory of management (Cyert and March 1963). Simon never declared agents to be irrational or crazy, simply unavoidably bounded by limits that they must face and operate within.

A curious matter here has been the occasional effort by more standard neoclassical economists to try to subsume Simon and his view into their worldview. Thus Stigler (1961) argued that Simon's view simply amounted to adding another variable to be optimized in the standard model, namely minimizing the costs of information. If full information is impossible due to infinite cost, then one estimates just the right amount of information to obtain. This sounds good on the surface, but it ignores the problem that people do not know what the full costs of information are. They may need to pursue a higher level activity: determining the costs of information. But that then implies yet another round of this: determining the costs of determining the costs of information, yet another ineluctable infinite regress as ultimately appears in Keynes's beauty contest (Conlisk 1996), yet another example of complexity undermining the ability to obtain knowledge.

Just as Stigler attempted to put Simon's ideas into a narrow box, so others since have attempted to do so as well, including many behavioral economists. But drawing on multiple approaches to complexity, Simon's understanding of the nature of the relationship between knowledge and complexity stands on its own as special and worthy of its continuing influence (Velupillai 2019).

3.6 Conclusions

We have reviewed issues related to the problem of knowledge in complex systems. While there are many competing definitions of complexity, we have identified two that have been most frequently used in economics: dynamic complexity and computational complexity. Each has its own sort of epistemological problems. Dynamic complexity is subject to such issues as the sensitive dependence on initial conditions of chaos theory, or the uncertainty due to fractal basin boundaries in stochastic nonlinear systems, or the pattern of phase transitions and self-organizing transformations that can occur in systems with interacting heterogeneous agents. Such problems imply that in effect only an infinite degree of precision of knowledge will allow one to fully understand the system, which is impossible.

In computationally complex systems the problem is more related to logic, the problems of infinite regress and undecidability associated with self-referencing in systems of Turing machines. This can manifest itself as the halting problem, something that can arise even for a computer attempting to precisely calculate even a dynamically complex system as for example the exact shape of the Mandelbrot set (Blum et al. 1998). A Turing machine cannot understand fully a system in which its own decisionmaking is too crucially a part. However, knowledge of such systems may be gained by other means.

These computational problems as well as those arising in nonlinear dynamical systems were key to Herbert Simon formulating his concept of bounded rationality. This was reinforced by his initiation of the idea of hierarchical complexity as well.

In the end, the serious epistemological problems associated with complex economic systems do imply that there exist serious bounds on the rationality of economic agents. These bounds take many forms, inability to understand the internal relations of a system, inability to fully know crucial parameter values, inability to identify critical thresholds or bifurcation points, inability to understand the interactions of agents, especially when these agents are thinking about how each other are thinking about each others' thinking. Infinite regress problems imply non-decidability and non-computability for hyper-rational Turing machine agents. Thus, economic agents must ultimately rely on arbitrary acts and decisions, even if those simply involve deciding what will be the bounds beyond which the agent will no longer attempt to solve the epistemological problem.

References

Albin PS, Foley DK (1998) Barriers and bounds to rationality: essays on economic complexity and dynamics in interactive systems. Princeton University Press, Princeton

Arthur WB, Durlauf SN, Lane DA (1997) Introduction. In: Brian Arthur W, Durlauf SN, Lane DA (eds) The economy as an evolving complex system II. Addison-Wesley, Reading, pp 1–14

Binmore K (1987) Modeling rational players, I. Econ Philos 3:9–55

Binmore K, Samuelson L (1999) Equilibrium selection and evolutionary drift. Rev Econ Stud 66:363–394

Blum L, Cucker F, Shub M, Smale S (1998) Complexity and real computation. Springer, New York

Brock WA, Hommes CH (1997) A rational route to randomness. Econometrica 65:1059–1095

Brock WA, Hommes CH (1998) Heterogeneous beliefs and routes to chaos in a simple asset pricing model. J Econ Dyn Control 22:1235–1274

Canning D (1992) Rationality, computability, and nash equilibrium. Econometrica 60:877–888

Chaitin GJ (1987) Algorithmic information theory. Cambridge University Press, Cambridge
Chen S-H (2003) Agent-based computational macroeconomics: a survey. In: Terano T, Deguchi H, Takadama K (eds) Meeting the challenge of social problems via agent-based simulation. Springer, New York, pp 141–170
Chen S-H (2005) Computational intelligence in economics and finance: carrying on the legacy of Herbert Simon. Inf Sci 170:121–131
Chen S-H (2012) Varieties of agent-based computational economics: a historicaal and an interdisciplinary perspective. J Econ Dyn Control 36:1–25
Chen S-H (2014) Neuroscience and agent-based computational economics. Int J Appl Behav Econ 3(2). 104018/ijabe.2014040102
Chen S-H (2016) Agent-based computational economics: how the idea originated and where it is going. Routledge, London
Chen S-H, Chie B-T (2008) Lottery markets design, micro-structure, and macro-behavior: an ACE approach. J Econ Behav Organ 67:463–480
Chen S-H, Kuo Y-F (2016) Herbert Simon and agent-based computational economics. In: Frantz R, March L (eds) Minds, models, and milieux: commemorating the centennial birth of Herbert Simon. Palgrave Macmillan, New York, pp 113–144
Chen S-H, Yeh C-H (1996) Genetic programming and the cobweb model. In: Angeline P (ed) Advances in genetic programming, vol 2. MIT Press, Cambridge, MA, pp 443–466
Chen S-H, Yeh CH (1997) Toward a computable approach to the efficient market hypothesis: an application of genetic programming. J Econ Dyn Control 21:1043–1063
Chen S-H, Yeh CH (2002) On the emergent properties of artificial stock markets: the efficient market hypothesis and the rational expectations hypothesis. J Econ Behav Organ 49:217–230
Church A (1936) A Note on the Entscheidungs problem. J Symb Logic 1:40–41, Correction 101–102
Conlisk J (1996) Why bounded rationality? J Econom Literat 34:1–64
Cyert RM, March JG (1963) A behavioral theory of the firm. Prentice-Hall, Englewood Cliffs
Day RH (1994) Complex economic dynamics, volume i: an introduction to dynamical systems and market mechanisms. MIT Press, Cambridge
Diaconis P, Freedman D (1986) On the consistency of bayes estimates. Ann Stat 14:1–26
Gode D, Sunder S (1993) Allocative efficiency of markets with zero intelligence traders: markets as a partial substitute for individual rationality. J Polit Econ 101:119–137
Hayek FA (1948) Individualism and economic order. University of Chicago Press, Chicago
Hayek FA (1967) The theory of complex phenomena. Studies in philosophy, politics and economics. Routledge & Kegan Paul, London, pp 22–42
Horgan J (1997) The end of science: facing the limits of knowledge in the twilight of the scientific age, paperback. Broadway Books, New York
Ijiri Y, Simon HA (1964) Business firm growth and size. Am Econ Rev 54:77–89
Israel G (2005) The science of complexity: epistemological problems and perspectives. Sci Context 18:1–31
Kao YF, Ragupathy V, Vela Velupillai K, Zambelli S (2012) Noncomputability, unpredictability, undecidability, and unsolvability in economic and finance theories. Complexity 18:51–55
Keynes JM (1936) The general theory of employment, interest and money. Macmillan, London
Kindleberger CP (1978) Manias, panics, and crashes. Basic Books, New York
Kleene SC (1967) Mathematical logic. Wiley, New York
Kolmogorov AN (1965) Combinatorial foundations of information theory and the calculus of probabilities. Russ Math Surv 38(4):29–40
Koppl R, Rosser JB, Jr (2002) All that i have to say has already crossed your mind. Metroeconomica 53:339–360
Landini S, Gallegati M, Rosser JB, Jr (2020) Consistency and incompleteness in general equilibrium theory. J Evol Econ 30:205–230
Lawson T (1997) Economics and reality. Routledge, London
Lewis AA (1985) On effectively computable realizations of choice functions. Math Soc Sci 10:43–80

Lipman BL (1991) How to decide how to decide how to…,: modeling limited rationality. Econometrica 59:1105–1125
Loasby BJ (1976) Choice, complexity and ignorance. Cambridge University Press, Cambridge
Lorenz EN (1963) Deterministic non-periodic flow. J Atmos Sci 20:130–141
Lorenz H-W (1992) Multiple attractors, complex basin boundaries, and transient motion in deterministic economic systems. In: Feichtinger G (ed) Dynamic economic models and optimal control. North-Holland, Amsterdam, pp 411–430
Markose SM (2005) Computability and evolutionary complexity: markets as complex adaptive systems. Econ J 115:F159–F192
Minsky HP (1972) Financial instability revisited: the economics of disaster. Reappraisal Federal Reserve Discount Mech 3:97–136
Mirowski P (2002) Machine dreams: economics becomes a cyborg science. Cambridge University Press, Cambridge
Mirowski P, Nik-Kah E (2017) Knowledge we have lost in information: the history of information in modern economics. Oxford University Press, New York
Nyarko Y (1991) Learning in mis-specified models and the possibility of cycles. J Econ Theory 55:416–427
Pryor FL (1995) Economic evolution and structure: the impact of complexity on the U.S. economic system. Cambridge University Press, New York
Richter MK, Wong KV (1999) Non-computability of competitive equilibrium. Econ Theor 14:1–28
Rissanen J (1989) Stochastic complexity in statistical inquiry. World Scientific, Singapore
Rosser JB, Sr (1936) Extensions of some theorems of Gödel and Church. J Symb Log 1:87–91
Rosser JB, Jr (1991) From catastrophe to chaos: a general theory of economic discontinuities. Kluwer, Boston
Rosser JB, Jr (1996) Chaos and rationality. In: Elliott E, Douglas Kiel L (eds) Chaos theory in the social sciences. University of Michigan Press, Ann Arbor, pp 199–212
Rosser JB, Jr (1998) Complex dynamics in new Keynesian and post Keynesian models. In: Rotheim RJ (ed) New Keynesian economics/post Keynesian alternatives. Routledge, London, pp 288–302
Rosser JB, Jr (1999) On the complexities of complex economic dynamics. J Econ Perspect 13(4):169–192
Rosser JB Jr (2004) Epistemological implications of economic complexity. Ann Japan Assoc Philos Sci 31(2):3–18
Rosser JB, Jr (2010) Constructivist logic and emergent evolution in economic complexity. In: Zambelli S (ed) Computability, constructive and behavioual economic dynamics: essays in honour of Kumaraswamy (Vela) Velupillai. Routlege, London, pp 184–197
Rosser JB, Jr (2012) On the foundations of mathematical economics. New Math Natl Comput 8:53–72
Rosser JB, Jr (2014) The foundations of economic complexity in behavioral rationality in heterogeneous expectations. J Econ Methodol 21:308–312
Rosser JB, Jr (2020a) Reflections on reflexivity and complexity. In: Dolfsma W, Wade Hands C, McMaster R (eds) History, methodology and identity for a 21st social economics. Routledge, London, pp 67–86
Rosser JB, Jr (2020b) The minsky moment and the revenge of entropy. Macroecon Dyn 24:7–23
Rosser JB Jr (2021) Foundations and applications of complexity economics. Springer Nature, Cham
Rosser JB Jr, Rosser MV (2015) Complexity and behavioral economics. Nonlinear Dyn Psychol Life Sci 19:67–92
Rosser JB Jr, Rosser MV, Gallegati M (2012) A Minsky-Kindleberger perspective on the financial crisis. J Econ Issues 45:449–458
Shackle GLS (1972) Epistemics and economics: a critique of economic doctrines. Cambridge University Press, Cambridge
Simon HA (1955) A behavioral model of rational choice. Quart J Econ 60:99–118
Simon HA (1957) Models of man: social and rational. Wiley, New York
Simon HA (1962) The architecture of complexity. Proc Am Philos Soc 106:467–482

Simon HA (1969) The sciences of the artificial. MIT Press, Cambridge
Smith A (1759) The theory of moral sentiments. Miller, Kincaid & Bell, London
Stigler GJ (1961) The economics of information. J Polit Econ 69:213–225
Stodder JP (1995) The evolution of complexity in primitive economies: theory. J Compar Econ 20:1–31
Turing AM (1937) Computability and λ-definability. J Symb Log 2:153–163
Velupillai K (2000) Computable economics. Oxford University Press, Oxford
Velupillai KV (2019) Classical behavioural finance theory. Rev Behav Econ 6:1–18
Zambelli S (2004) Production of ideas by means of ideas: turing machine metaphor. Metroeconomica 55:155–179
Zeeman EC (1974) On the unstable behavior of the stock exchanges. J Math Econ 1:39–44

Chapter 4
Agent-Based Computational Economics: Overview and Brief History

Leigh Tesfatsion

Abstract Scientists and engineers seek to understand how real-world systems work and could work better. Any modeling method devised for such purposes must simplify reality. Ideally, however, the modeling method should be flexible as well as logically rigorous; it should permit model simplifications to be appropriately tailored for the specific purpose at hand. Flexibility and logical rigor have been the two key goals motivating the development of *Agent-based Computational Economics (ACE)*, a completely agent-based modeling method characterized by seven specific modeling principles. This perspective provides an overview of ACE, a brief history of its development, and its role within a broader spectrum of experiment-based modeling methods.

Keywords Completely agent-based modeling (c-ABM) · Agent-based computational economics (ACE) · Experiment-based modeling methods

JEL Codes C6 · C7 · D9 · E7

4.1 Introduction

The term *Agent-Based Modeling (ABM)* refers to a class of modeling methods designed for the study of systems whose dynamics are driven by successive interactions among heterogeneous entities. Such systems range from the particle systems

This ACE perspective, based in part on Tesfatsion (2017, 2021b, 2022a), is an invited chapter for: R. Venkatachalam (Ed.), 2022. *Artificial Intelligence, Learning and Computation in Economics and Finance*, Springer, to appear. A preprint version of this ACE perspective is posted at the ISU Digital Repository as Economics Working Paper #21004; see Tesfatsion (2022j).

L. Tesfatsion (✉)
Research Professor and Professor Emerita of Economics, Courtesy Research Professor of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011-1054, USA
e-mail: tesfatsi@iastate.edu
URL: https://www2.econ.iastate.edu/tesfatsi/

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_4

studied in physics to the coupled human and natural systems studied in socio-ecology. Consequently, the pathways leading to the development of ABM cannot be depicted as a tree, or even as a gnarly bush, but instead must be envisioned as a forest of diverse trees supported by a complex interconnected network of roots.

Many previous authors have ably explored the various origins and meanings of ABM; see, for example, Arthur (2021), Axtell and Farmer (2022), Chen (2002, 2012, 2016), Epstein (2006), Epstein and Axtell (1996), Gallegati (2018), Kirman (2011), Railsback and Grimm (2019), and Wilensky and Rand (2015). The purpose of this perspective is much more modest in scope: namely, to discuss the origin and development of one particular variant of ABM called *Agent-based Computational Economics (ACE)*.

I chose the name "ACE" in Ames, Iowa, in August 1996, following my participation in a satellite meeting held at the end of the Second International Conference on Computing in Economics and Finance (CEF1996, Geneva, Switzerland, June 26–28, 1996). A key purpose of this satellite meeting was to discuss how ABM could be promoted within the economics profession. I came away from this meeting determined to develop a website devoted to this objective, and I needed a website name that would clearly convey to other economists that this modeling method differed in essential regards from then-standard economic modeling methods.

Through the years, however, I have come to realize that my conception of ACE modeling also differs in essential regards from other variants of ABM. For example, I have always considered an ACE agent to be a *software* entity within a *computationally-constructed world*, characterized at each instant by its current *state* (*data, attributes, and/or methods*). Moreover, in keeping with the standard dictionary definition for *agent*, I have always required an ACE agent to be capable of affecting the trajectory of outcomes for its world. Subject to these conditions, I have permitted ACE agents to represent a broad range of entities: e.g., individual life-forms, social groupings, institutions, and/or physical phenomena. However, I have always insisted that the resulting ACE model be *completely* agent-based in the following sense: Given initial modeler-specified agent states, *all model dynamics are driven by agent interactions*.

Although I have consistently viewed these modeling principles to be necessary underpinnings for any ACE model, clearly these principles are not specific to economic systems. Rather, together with additional supporting principles, they characterize a completely agent-based variant of ABM that I now refer to as *completely Agent-Based Modeling (c-ABM)*.

Section 4.2 of this chapter provides an axiomatic characterization of c-ABM, expressed in terms of seven specific modeling principles. The potential usefulness of c-ABM for the study of general real-world systems is considered in Sect. 4.3. ACE is defined in Sect. 4.4 to be the specialization of c-ABM to economic systems. The ability of ACE agents to embody wide ranges of rationality and different forms of stochasticity is addressed in Sects. 4.5 and 4.6; and four current ACE research directions, delineated by objective, are described in Sect. 4.7. The history of ACE is briefly outlined in Sect. 4.8, documented in part by archived copies of ACE news notes that I distributed from 1997 through 2017.

Finally, the concluding Sect. 4.9 considers two intriguing directions for future research. First, the actions undertaken by the constituent agents of a c-ABM world can be based on non-constructive beliefs ("leaps of faith") as well as on constructive beliefs resulting from directly observed or experienced world events. This capability could facilitate the study of real-world systems that are complex blends of physical and social processes. Second, human-subject studies and c-ABM constitute the two polar end-points for a promising spectrum of hybrid human/agent experiment-based modeling methods in need of more systematic exploration.

4.2 Completely Agent-Based Modeling (c-ABM)

Roughly defined, *completely Agent-Based Modeling (c-ABM)* is the computational modeling of processes as open-ended dynamic systems of interacting agents. Here an "agent" for a system is broadly construed (in a traditional dictionary sense) to be any entity capable of affecting the trajectory of outcomes for this system. Agents can thus range from sophisticated strategic decision-making entities (e.g., "humans") to physical phenomena with no cognitive function (e.g., "weather").

An axiomatic characterization of c-ABM is given below in terms of seven modeling principles. These principles are not strictly independent of each other. However, each principle stresses a distinct c-ABM feature, as indicated by its caption. Together, these seven modeling principles reflect the primary goal of many agent-based modelers: namely, to be able to study real-world dynamic systems as historically unfolding events, driven by agent interactions.

(MP1) Agent Definition: An *agent* is a software entity within a computationally-constructed world that can affect world outcomes through expressed actions.
(MP2) Agent Scope: Agents can represent a broad range of entities, e.g., individual life-forms, social groupings, institutions, and/or physical phenomena.
(MP3) Agent Local Constructivity: An intended action of an agent at a given instant is determined by the agent's *state* (*data, attributes, and/or methods*) at this instant.
(MP4) Agent Autonomy: All *agent interactions* (*expressed agent actions*) at a given instant are determined by the ensemble of agent states at this instant.
(MP5) System Constructivity: The state of the world at a given instant is determined by the ensemble of agent states at this instant.
(MP6) System Historicity: Given an initial ensemble of agent states, any subsequent *world event* (*change in agent states*) is induced by prior and/or concurrent agent interactions.
(MP7) Modeler as Culture-Dish Experimenter: The role of the modeler is limited to the configuration and setting of initial agent states, and to the non-perturbational observation, analysis, and reporting of world outcomes.

The first six modeling principles (MP1)–(MP6) characterize an agent-based model in initial-value state-space form.[1] More precisely, they specify how an ensemble of agent states dynamically evolves, starting from an initially given ensemble of agent states, where each agent state consists of data, attributes, and/or methods. This dynamic evolution is required to exhibit four essential real-world properties: namely, agent local constructivity, agent autonomy, system constructivity, and system historicity. The seventh modeling principle (MP7) limits the role of the modeler in the modeling process to the configuration and setting of initial agent states.

Considered as a whole, the seven modeling principles (MP1)–(MP7) thus characterize a completely agent-based model as a computational laboratory permitting the exploration of a computationally-constructed world. This exploration process is analogous to biological experimentation with cultures in Petri dishes. The modeler configures and sets initial conditions for the world. The modeler then steps back, assuming the role of pure observer, as subsequent world events are driven solely by the interactions of the world's constituent entities.

4.3 C-ABM: A Mathematics for the Real World?

Modeling flexibility and logical rigor have been the two key goals motivating my development of the seven c-ABM modeling principles (MP1)–(MP7) presented in Sect. 4.2. First, having entered into economics at the graduate level, following an undergraduate degree in history, I wanted to be able to model and study real-world economic systems as historical processes whose "human" participants were able to "breathe." Second, as a mathematical economist trained at the University of Minnesota, I wanted this mode of economic modeling to be clearly delineated as a rigorous flexible alternative to the highly constrained modes of economic modeling I had been encountering in economic textbooks and journals.

Tesfatsion (2017, 2021b) provide careful justification and illustrative support for the contention that c-ABM provides a flexible and rigorous modeling method for the study of real-world economic systems involving complex intertwined social and physical processes. An interesting speculative question is the extent to which c-ABM provides a useful modeling method for the study of real-world systems *in general*.

Any system modeled in accordance with (MP1)–(MP7) is an open-ended dynamic system of interacting agents, each characterized by its own state (data, attributes, and/or methods). These agents can represent any entity capable of affecting the trajectory of system outcomes: e.g., individual life-forms, social groupings, institutions, and/or physical phenomena. The interactions of these agents induce all dynamics (state changes) for the modeled system, starting from initial agent states configured and set by the modeler. As a result of these interactions:

[1] An *initial-value* state-space model is a state-space model for a dynamic system S that runs forward through time, commencing at some specified start-time t^o, with all boundary conditions taking the form of constraints on the state of S at the start-time t^o.

– each agent experiences "time" locally, as an unfolding sequence of events;
– the dimension and content of agent states can change;
– agents can subsume other agents as components;
– agents can break apart into smaller component agents;
– new agents can be created;
– existing agents can be destroyed.

Examples of *state-changes* for real-world agents include: changes in sensed surroundings; changes in recorded observations; changes in physical attributes; changes in beliefs; and belief-induced changes in action rules. Examples of real-world agent *subsumption* include: the formation of molecules through atomic bonding; the transition from prokaryotic to eukaryotic forms of organisms; the parasitism of one organism by another; the hiring of employees by corporate firms; the acquisition of new members by existing organizations; and the merger of organizations.

Examples of real-world agent *creation and destruction* include: volcanic eruptions; natural birth and death; the invention and obsolescence of products; and the establishment and disbanding of organizations. Creation and destruction events for populations of agents can be computationally modeled by means of evolutionary algorithms taking various forms.

Note, in particular, that models adhering to (MP1)–(MP7) permit the study of real-world systems that evolve from initial conditions with:

– **no fixed "space"** apart from persistent spatial agents (if any) that modelers initially configure;
– **no fixed "time process"** apart from persistent event-scheduler agents (if any) that modelers initially configure;
– **no fixed "physical laws"** apart from persistent agent methods (if any) that modelers initially configure.

The ability to model real-world systems without having to presuppose a fixed externally given "space" or "time process" permits the study of open perplexing questions in physics regarding the existence (or not) of these concepts as fundamental fixed aspects of the physical universe.

Persistent agent methods that a researcher might want to initially configure for a modeled real-world system include methods that support self-organization and natural selection processes. These types of processes appear to be a basic driver of real-world agent interactions at all levels of agent encapsulation that humans can currently perceive. An interesting question is whether they also drive agent interactions at levels beyond current human perception, such as at a quantum level.

Finally, c-ABM permits the "thickly constructive" modeling of real-world systems in the following sense: Given initial agent states, to an external observer the model might appear to consist of successive *changes* in agent states constructively determined by successive agent interactions. In actuality, these successive agent interactions are determined by successive agent states whose evolution can entail non-constructive "leaps of faith."

A more precise characterization of c-ABM as a *thickly-constructive* modeling method is as follows. By definition, the state of an agent at a given instant consists of data, attributes, and/or methods. By agent local constructivity and autonomy, all agent interactions at a given instant are determined by the ensemble of agent states at this instant. By system historicity, any world event (change in agent states) at a given instant is induced by prior or concurrent agent interactions. However, an agent's state at a given instant can include acquired or evolved attributes taking the form of *non-constructive beliefs*, i.e., beliefs that assign truth values to propositions that are not constructively decidable. Consequently, non-constructive agent beliefs ("leaps of faith") at a given instant can affect future world events.

Models satisfying the seven c-ABM modeling principles (MP1)–(MP7) thus permit non-constructive agent beliefs to function as possible causal factors for Carlo Rovelli's "world...of events, not things" (Rovelli 2018, Chap. 6), Gilbert Ryle's "ghost in the machine" (Ryle 1949, pp. 11–24), and Lee Smolin's "seers" (Smolin 2007, Part IV.18).

4.4 Agent-Based Computational Economics

Agent-based Computational Economics (ACE) is the specialization of *completely Agent-Based Modeling (c-ABM)* to economic systems. More precisely, ACE is the modeling of *economic* systems in accordance with the seven c-ABM modeling principles (MP1)–(MP7) presented in Sect. 4.2.

Each ACE model must therefore be an initial-value state-space economic model satisfying the agent definition (MP1), the agent scope requirement (MP2), and the five additional requirements (MP3)–(MP7): namely, agent local constructivity; agent autonomy; system constructivity; system historicity; and modeler as culture-dish experimenter. As detailed in Tesfatsion (2017), ACE thus permits economists to study real-world economies as open-ended locally-constructive sequential games.

Modern economic theory also relies heavily on state-space models. However, these models typically incorporate modeler-imposed rationality, optimality, and/or equilibrium conditions that could not (or would not) be met by locally-constructive and autonomous agents interacting within economic systems that satisfy system constructivity and system historicity. For example, strong-form rational expectations assumptions require the ex ante expectations of decision-makers to be consistent with ex post model outcomes. The determination of a rational expectations solution is therefore a global fixed-point problem that requires the simultaneous consideration of all modeled decision periods without regard for local constructivity, autonomy, and historical process constraints.

In contrast, ACE permits the *open-ended* dynamic modeling of economic systems *without external imposition* of rationality, optimality, or equilibrium conditions. ACE models can therefore be used to conduct systematic investigations of these conditions as testable prior hypotheses. This capability fundamentally distinguishes ACE from all currently standard dynamic economic modeling methods.

Finally, the requirement that ACE models satisfy the seven c-ABM modeling principles (MP1)–(MP7) permits ACE to be distinguished more clearly and carefully from other variants of agent-based modeling (Chen 2016, Chaps. 1–2), and from important related modeling methods such as microsimulation (Richiardi 2013), system dynamics (Rahmandad and Sterman 2008), and econophysics (Chen and Li 2012).

4.5 ACE Agent Rationality

For ACE researchers, as for economists in general, the modeling of decision-makers is a primary concern. Consequently, it is important to correct a major misconception still being expressed by some economic commentators uninformed about the powerful capabilities of modern software: namely, the misconception that ACE decision-making agents cannot be as rational (or irrational) as real-world decision-makers.

To the contrary, the constraints on agent decision-making implied by the seven c-ABM modeling principles (MP1)–(MP7) are constraints inherent in every real-world dynamic system. As demonstrated concretely in Sinitskaya and Tesfatsion (2015), the methods used by ACE decision-making agents can range from simple behavioral rules to sophisticated anticipatory learning algorithms for the approximate achievement of intertemporal objectives.

Extensive annotated pointers to introductory materials on the implementation of learning and decision methods for ACE agents can be accessed at the ACE learning research repository (Tesfatsion 2022e). The learning methods covered in these materials include:

- *Reactive reinforcement learning.* Roth-Erev reactive reinforcement learning, ... ;
- *Belief-based learning.* Fictitious play, Camerer/Ho EWA algorithm, ... ;
- *Anticipatory learning.* Q-learning, adaptive dynamic programming, ... ;
- *Evolutionary learning.* Genetic algorithms, genetic programming, ... ;
- *Connectionist learning.* Associative memory learning, artificial neural network (ANN) learning, deep learning using ANNs with multiple hidden layers, ...

4.6 ACE Agent Stochasticity

Stochastic aspects can easily be represented within ACE models. ACE agent data can include past or run-time realizations for real-world random events, ACE agent attributes can include beliefs based on probabilistic assessments, and ACE agent methods can include *Pseudo-Random Number Generators (PRNGs)*.

A PRNG is a deterministic algorithm A, initialized by a seed value s, able to generate a sequence $A(s)$ of numbers with the following property: Over some finite initial length $L(s)$, the sequence $A(s)$ closely mimics the properties of a random number sequence. The typical length of $L(s)$ calculated across admissible seed values s is a key metric used to evaluate the performance quality of a PRNG A.

PRNGs can be included among the methods of ACE decision-making agents, thus permitting these agents to "randomize" their behaviors. For example, an ACE decision-making agent can use PRNGs to choose among equally preferred actions or action delays, to construct mixed strategies in game situations to avoid exploitable predictability, and/or to induce perturbations in action routines in order to explore new action possibilities.

PRNGs can also be included among the methods of other types of ACE agents, such as ACE physical or biological agents, in order to model stochastic phenomena external to ACE decision-making agents. For example, an ACE weather agent can use a PRNG to generate a weather pattern for its computational world during a simulated time-interval T that affects the actions expressed by ACE decision-making agents during T.

An additional important point is that ACE agents are *encapsulated* in the following sense: The internal data, attributes, and/or methods of each ACE agent A can be partially or completely hidden from any other ACE agent B, either by the deliberate choice of agent A, or by initial modeler specification. Thus, ACE agents can be unpredictable to one another even if they make no use of random event realizations, probabilistic assessments, or PRNGs.

Finally, the seven c-ABM modeling principles (MP1)–(MP7), considered as a whole, require ACE models to be *stochastically complete* in the following sense: If an ACE modeler desires to include a simulated stochastic shock process within their computationally-constructed world, the source (originating point) and sinks (impact points) for this shock process must be explicitly represented as agents that reside and interact within this world. Stochastic completeness thus encourages ACE modelers to think carefully about the intended empirical referents for any simulated stochastic shock processes. This, in turn, should help to reduce or eliminate reliance on *ad hoc* external shock terms as the sources of dynamic persistence and the drivers of dynamic interactions.

4.7 ACE Research Objectives

Current ACE research divides roughly into four branches, each corresponding to a different objective.

One primary objective is **understanding the appearance and persistence of empirical regularities**. Examples include adherence to social norms, socially accepted monies, widely instituted market protocols, business cycles, persistent wealth inequality, and the common adoption and use of technological innovations.

An ACE model capable of generating an empirical regularity based on empirically-credible agent specifications provides a candidate explanation for this regularity. As discussed more carefully by LeBaron and Tesfatsion (2008) and Tesfatsion (2017, 2022d), the empirical validation of agent specifications should ideally encompass four distinct aspects: (i) *Input Validation:* Validation of initially specified agent data and attributes; (ii) *Process Validation:* Validation of initially specified agent methods; (iii) *Descriptive Output Validation:* In-sample model fitting; and (iv) *Predictive Output Validation:* Out-of-sample model forecasting.

A second primary objective is **normative design**. How can ACE models facilitate the design of structures, institutions, policies, and/or regulations intended to improve the performance of economic systems? The ACE approach to normative design is akin to filling a bucket with water to determine if it leaks. An ACE researcher computationally constructs a world capturing salient aspects of an economic system operating under a proposed design. The researcher identifies a range of initial agent state specifications of interest, including seed values for agent PRNG methods. For each such specification the researcher permits the world to develop forward, driven solely by agent interactions. Recorded world outcomes are then used to evaluate design performance.

A critical issue for ACE normative design studies is the extent to which outcomes resulting under a tested design are efficient, fair, and orderly, despite possible attempts by ACE decision-making agents to game the design for personal advantage. A related issue is a cautionary concern for adverse unintended consequences. *Optimal* design might not be achievable, especially for large complex systems; but ACE modeling can facilitate *robust* design for increased system reliability and resiliency, a goal that is both feasible and highly desirable.

A third primary objective is **qualitative insight and theory generation**. How can ACE modeling be used to study the *potential* future behavior of an economic system? A quintessential example of this line of research is an old but still unresolved concern of economists such as Adam Smith (1723–1790), Ludwig von Mises (1881–1973), John Maynard Keynes (1883–1946), Joseph Schumpeter (1883–1950), and Friedrich von Hayek (1899–1992): namely, what are the self-organizing capabilities of decentralized market economies?

Ideally, what is needed for this objective is the *phase portrait* of the economic system, i.e., a representation of the system's potential state trajectories starting from each possible initial system state. This phase portrait would help to clarify which regions of the system's state space are credibly reachable, hence of empirical interest, and which are not. It would also reveal the possible existence of equilibrium state trajectories E, however "equilibrium" is defined. Finally, it would reveal the *basin of attraction* for any such E, that is, the (possibly empty) subset of system states which, if reached, would result in progression to E.

An ACE modeling of an economic system permits the modeler to conduct batched model runs, starting from multiple initial agent-state specifications. The modeler can thus generate a rough approximation of the system's phase portrait.

A fourth primary objective is **method/tool advancement**. How best to provide ACE researchers with the methods and tools they need to undertake theoretical studies of dynamic economic systems through systematic sensitivity studies, and to examine the compatibility of sensitivity-generated theories with real-world data? ACE researchers are exploring a variety of ways to address this objective ranging from the careful consideration of methodological principles to the practical development of programming, verification, empirical validation, and visualization tools.

4.8 Brief History of ACE

I first encountered agent-based modeling in a delightful 1983 *Scientific American* essay (Hofstadter 1983) celebrating Bob Axelrod's work on *Iterated Prisoner's Dilemma (IPD)* tournaments (Axelrod 1984).[2] Axelrod's key idea was first to specify an initial population of computer programs, each implementing an IPD strategy, and then to let these programs engage in repeated round-robin play of PD games with or without evolution of their initially programmed strategies. The goal was to see under what conditions, and to what extent, cooperative play might be induced.

Two aspects of Axelrod's tournaments stood out for me in comparison with standard economic modeling approaches at the time. First, even in deterministic form, the tournaments involved sufficiently complex interactions that it was difficult to deduce long-run outcomes from initial conditions. Thus, as in real-world biological experiments with cultures in Petri dishes, researchers could be genuinely surprised by tournament outcomes. Second, in repeated play, Axelrod's agents (computer programs) exhibited induced "social" behaviors with interesting "life-like" characteristics, such as trust, deception, reciprocity, and stance towards strangers.

In the mid-1980s I was heavily involved in the development of *adaptive computation* methods, i.e., flexible computational solution methods able to adapt to the problem at hand rather than requiring the problem to be adapted to the method. It thus took me some time to redirect my research towards an exploration of Axelrod's intriguing agent-based approach for the flexible modeling of economic systems.

Indeed, my first "agent-based" work was a 1991 adaptive computation paper (Kalaba and Tesfatsion 1991) co-authored with applied mathematician Bob Kalaba. In this paper we develop an *adaptive homotopy* solution method able, in runtime, to detect and avoid regions of the solution space where calculations are ill-conditioned due to nearby singularities or bifurcation points. This adaptive capability is achieved by replacing the standard homotopy continuation parameter, moving in a *pre-set* manner from 0 to 1 *along the real line*, with a "smart agent" able to construct and traverse an *adaptively-determined* path from 0+0i to 1+0i *in the complex plane*. The smart agent decides the direction and length of its next step, given its current state, by solving a multi-criteria optimization problem requiring a trade-off between two criteria: (i) maintain a short path length from 0+0i to 1+0i; and (ii) avoid regions in the complex plane where ill-conditioning of calculations is detected.

During the early-to-mid 1990s I increasingly participated in ABM-related conference panel sessions. This participation included: the Artificial Life III Conference sponsored by the Santa Fe Institute (Sweeney Center, Santa Fe, New Mex-

[2] This essay was brought to my attention circa 1985 by Bob Rider, a Ph.D. student in the Department of Economics at the University of Southern California with an interest in game theory.

ico, June 15–19, 1992); a session at the Economic Science Association Meeting (Tucson, Arizona, October 21–23, 1993); a session at the North American Summer Meeting of the Econometric Society (Université Laval, Quebec City, June 24–28, 1994); the First International Conference on Computing in Economics and Finance (CEF1995, University of Texas, Austin, May 21–24, 1995); the First Economic Artificial Life Conference (Santa Fe Institute, Santa Fe, New Mexico, May 26–29, 1995); an American Economic Association panel session at the Annual Meeting of the Allied Social Science Associations (ASSA, San Francisco, CA, January 5–7, 1996); the UCLA Economic Simulation Conference (University of California, Los Angeles, February 9, 1996); and the Fifth Annual Conference on Evolutionary Programming (San Diego, California, February 29–March 2, 1996).

However, the most pivotal meeting for me, personally, was an informal "agent-based economics" meeting I organized, held immediately after the formal close of the Second International Conference on Computing in Economics and Finance (CEF1996, Geneva, Switzerland, June 26–28, 1996).[3] A key agenda item for this informal meeting was to consider how agent-based modeling might best be promoted to the economics profession at large. I left this meeting determined to develop a website devoted to this objective.

Exploiting the brand-new availability of web browsers, in particular Netscape Navigator,[4] I began my *Agent-Based Economics (ABE)* website in late July of 1996 on an Iowa State University (ISU) server. In addition, with important input from Rob Axtell, I supplemented the ABE website with an ABE mailing list to be used for the distribution of occasional news notes.

However, microeconomists at ISU and elsewhere—Herman Quirmbach in particular—soon convinced me that ABE was a poor name-choice. They predicted that economic theorists would be dismissive of this "new" agent-based modeling approach on the grounds that standard micro-founded economic models were already "agent-based" since they modeled the optimizing behaviors of individual consumers and/or firms. Crucial additional "agent-based" requirements (e.g., agent autonomy and system historicity) would be ignored. Consequently, as documented in the February 1997 ACE news notes (Tesfatsion 2022b),[5] I changed the names of my website and mailing list to *Agent-based Computational Economics (ACE)* in August 1996

[3] As indicated by a preserved sign-up sheet, the participants in this informal meeting were: Rob Axtell; Ann Bell; Chris Birchenhall; Kai Brandt; Thomas Brenner; Charlotte Bruun; Shu-Heng Chen; Michael Gordy; Sergei Guriev; Armin Haas; Esther Hauk; Gillioz Jean-Blaise; Alan Kirman; Bob Marks; Christian Rieck; Ernesto Somma; Leigh Tesfatsion; and Nick Vriend.

[4] Netscape Communications Corporation, founded in April 1994 by Marc Andreessen and James H. Clark, released Netscape Navigator in November 1994 as freely downloadable software. Netscape Navigator, a successor of Mosaic (co-developed by Andreesen), was among the first browser products released in support of the mid-1990s consumer Internet revolution.

[5] The earliest distributed ABM/ACE news notes were not saved in retrievable form; the online posted ACE news notes (Tesfatsion 2022b) begin in February 1997. The formatting of these online ACE news notes is ancient by browser standards. Although some formatting commands used in these news notes no longer compile properly using modern browsers, the news notes have been left in their originally posted form in order to preserve their historical authenticity.

to stress computational modeling as one feature distinguishing the proposed agent-based modeling approach from then-standard economic modeling approaches.[6]

This 1996 name-change from ABE to ACE turned out to be fortuitous. It immediately connected the ACE modeling method to seminal work on "computational economics" being undertaken by Ken Judd and other participants in the Society for Computational Economics (SCE), founded in 1995. ACE was soon formally named an SCE Special Interest Group, thus permitting its consideration for panel session allotment at annual SCE meetings.[7]

A major ACE landmark occurred in the summer of 1997. As documented in my ACE news notes (Tesfatsion 2022b) distributed between February and May of 1997, Program Chair Ken Judd invited Blake LeBaron and myself to organize two contributed-paper sessions on ACE for the Third International Conference on Computing in Economics and Finance (CEF1997) to be held in July 1997 at Stanford University, plus a post-meeting satellite session devoted entirely to ACE topics.

A second major ACE landmark occurred in 1998: I was invited to guest-edit three special journal issues on ACE, one for the *Journal of Economic Dynamics and Control (JEDC)* (Tesfatsion 2001a), another for *Computational Economics (CE)* (Tesfatsion 2001b), and a third for the *IEEE Transactions on Evolutionary Computation (IEEE TEC)* (Tesfatsion 2001c). As documented in my September 1998 ACE news notes (Tesfatsion 2022b), prospective authors for the JEDC special issue were asked to submit papers that addressed an issue of economic importance from an agent-based perspective. Prospective authors for the CE and IEEE TEC special issues were asked to submit papers with a strong agent-based computational component that addressed evolutionary economics issues.

These three ACE special issues all appeared in 2001. The research reported in these special issues demonstrated how ACE modeling permitted interesting ground-breaking extensions of then-standard economic modeling capabilities.

For example, Chen and Yeh (2001) develop an agent-based stock market model consisting of a collection of stock market traders together with a 'business school'. Each business faculty member at a given instant represents a particular 'school of thought' regarding the best stock market forecasting model. The comparative performance of these various forecasting models is regularly tested in a social review process (e.g., competition for publication in refereed journals), modeled via genetic programming. The business faculty use these test results to revise their models. A trader that takes time-off from trading to attend a particular faculty-member's course gains access to the forecasting model taught by this faculty member and tests whether this model outperforms his own currently-used model. If so, the trader replaces his currently-used model with the faculty member's model and returns to market trading. The stock market traders thus evolve their forecasting models using a combination of

[6] Specifically, to reflect the name change from ABE to ACE, the website URL address was changed from http://www.econ.iastate.edu/tesfatsi/abe.htm to http://www.econ.iastate.edu/tesfatsi/ace.htm, and the mailing list address was changed from abelist@iastate.edu to acenewslist@iastate.edu.

[7] The annual SCE meeting is officially referred to as the *International Conference on Computing in Economics and Finance (CEF)*.

individual learning (faculty course attendance decisions) and social learning (model replacement decisions).

As a second example, Tesfatsion (2001d) develops an agent-based labor market model with endogenous worker-employer matching, implemented by a Gale-Shapley deferred acceptance mechanism. To implement this mechanism, the workers and employers must exchange messages with each other at event-triggered instances regarding the receipt, acceptance, and refusal of work offers. During each labor market round, workers direct work offers to their most preferred employers; and employers accept work offers from their most preferred workers, refusing the rest. Once matched, a worker and employer engage in a work-site interaction modeled as a prisoner's dilemma game. The outcomes of these games in each labor market round affect worker and employer match preferences, hence who receives work offers and whose work offers are accepted or refused in the following round. This agent-based modeling of a labor market thus blends matching theory with game theory.

A third major ACE landmark occurred in 2005. Ken Arrow and Mike Intriligator, general editors for the North Holland (Elsevier) Handbooks in Economics Series, invited Ken Judd and myself to edit an ACE handbook volume for this series. Potential lead authors, with co-authors of their own choosing, were invited to submit chapters on topics of interest to ACE researchers.

Following a careful refereeing process, sixteen research chapters, seven perspective essays, and a resource guide for social science newcomers to agent-based modeling were accepted for the ACE handbook volume. The topic areas covered in the research chapters included: learning methods for economic agents; agent-based models and human-subject experiments; network formation among economic agents; agent-based computational finance; agent-based industrial organization; agent-based political economy; agent-based socio-economic modeling; agent-based software platforms for market design evaluation; and automated markets with trading agents. The ACE handbook volume (Tesfatsion and Judd 2006) was published in 2006.

As documented at the ACE website (Tesfatsion 2022a), research making use of agent-based modeling for the study of economic systems has greatly expanded since 2006. This research is now appearing in a variety of journals with a welcoming inclusive methodological stance.[8] Research areas include: auction markets; automated markets; business and management; coupled economic and ecological systems; development economics; economic policy; energy economics; evolution of institutions and social norms; experiments with real and computational agents; financial economics; industrial organization; labor economics; learning and the embodied mind; macroeconomics; network formation and evolution; organizations; path dependence and lock-in effects; political economy; and technological innovation.

Another welcome development, stressed in recent reviews (Arthur 2021; Axtell and Farmer 2022; Tesfatsion 2017), is that ACE researchers are increasingly focus-

---

[8] These welcoming economic journals include: *Computational Economics*; *International J. of Microsimulation*; *J. of Economic Behavior and Organization*; *J. of Economic Dynamics and Control*; *J. of Economic Interaction and Coordination*; and *J. of Evolutionary Economics*. For a more extensive linked listing of welcoming journals, including finance and game theory journals, see Tesfatsion (2022c).

ing on real-world applications in addition to conceptual advances. For example, as extensively documented at the research repositories (Tesfatsion 2022f, g, h) and in the survey articles (Dawid and Delli Gatti 2018; Dosi and Roventini 2019; Lori and Porter 2018; Tesfatsion 2018), three fast-growing application areas for ACE researchers are macroeconomic policy, financial economics, and electric power markets.

As a final note of optimism, consider the following: The new design proposed for centrally-managed wholesale power markets in the Wiley/IEEE Press book (Tesfatsion 2021a) was developed and tested by means of an open-source ACE platform (Battula and Tesfatsion 2020) that implements salient aspects of actual U.S. centrally-managed wholesale power markets. This use of an ACE platform did not elicit any negative comments from the editor or anonymous referees. Indeed, based on extensive refereeing for power system journals, my assessment is that agent-based computational platforms are now commonly used to model and study the daunting complexity of modern power system operations. Surely many real-world economic systems are at least as complex as real-world power systems.

4.9 Concluding Remarks

As detailed in previous sections, *Agent-based Computational Economics (ACE)* is a specialization of *completely Agent-Based Modeling (c-ABM)* to economic systems. In turn, c-ABM is a variant of ABM that is axiomatically characterized by the seven specific modeling principles (MP1)–(MP7) presented in Sect. 4.2.

Roughly summarized, any model satisfying (MP1)–(MP7) is a computationally-expressed initial-value state-space model consisting of a collection of agents (software entities), each characterized at a given instant by its current state (data, attributes, and/or methods). Given initial agent states, any subsequent event (change in agent states) is induced by prior and/or concurrent agent interactions. The role of the modeler is limited to the configuration and setting of initial agent states, and to the non-perturbational observation, analysis, and reporting of model outcomes.

Sections 4.2—4.3 provide support for the supposition that c-ABM provides the "right mathematics" for the study of general real-world systems, e.g., systems involving complex intertwined social and physical processes, or even purely physical processes. Viewing c-ABM as a form of "mathematics" is apt in the following sense:

- A ***classical mathematician*** specifies a model $\boldsymbol{M}$ in equation form that embodies structural assumptions regarding the relationship between $\boldsymbol{M}$'s externally-determined inputs $\boldsymbol{x}$ and the (possibly empty) set of resulting model-determined outputs (solutions) $\boldsymbol{v}$, where $\boldsymbol{x}$ is restricted to lie in some admissible input-space $\boldsymbol{X}$. The classical mathematician then *proves theorems* for $\boldsymbol{M}$ in two forms:
  - **Necessity of A for B** *(if B, then A)*: If a solution $\boldsymbol{v}'$ for $\boldsymbol{M}$ satisfies property $\boldsymbol{P}'$, then the input $\boldsymbol{x}'$ for $\boldsymbol{M}$ must lie in $\boldsymbol{X}' \subseteq \boldsymbol{X}$;
  - **Sufficiency of A for B** *(if A, then B)*: If the input $\boldsymbol{x}'$ lies in $\boldsymbol{X}' \subseteq \boldsymbol{X}$, then any corresponding solution $\boldsymbol{v}'$ for $\boldsymbol{M}$ must satisfy property $\boldsymbol{P}'$.

- A ***c-ABM modeler*** specifies a model $\boldsymbol{CM}$ in computational (software) form that embodies structural assumptions regarding the relationship between $\boldsymbol{CM}$'s externally-determined inputs $\boldsymbol{cx}$ (initial agent states) and the (possibly empty) set of resulting model-determined outputs (agent state trajectories) $\boldsymbol{cv}$, where $\boldsymbol{cx}$ is restricted to lie in some admissible input-space $\boldsymbol{CX}$. The c-ABM modeler then *implements an experimental study* for $\boldsymbol{CM}$ taking the following form: For each input $\boldsymbol{cx}'$ in a specified *finite* subset $\boldsymbol{CX}'$ of the input-space $\boldsymbol{CX}$, what properties are exhibited by any resulting computationally-generated output $\boldsymbol{cv}'$?

Critics might argue that a classical mathematician typically establishes necessity and sufficiency theorems (input $\leftrightarrow$ output relationships) for an analytically-expressed model $\boldsymbol{M}$ over *infinite* input subspaces $\boldsymbol{X}'$. In contrast, a c-ABM modeler can only establish sufficiency "examples" (input $\rightarrow$ output relationships) for a computationally-expressed model $\boldsymbol{CM}$ over *finite* input subsets $\boldsymbol{CX}'$. This criticism can be countered in two ways.

First, as eloquently argued by (Judd 2006, Sect. 4, p. 886), a "theorem" is simply a collection of "examples." The *relevance and robustness* of a collection of "examples" is surely more important than the *number* of these "examples."

Second, in classical mathematics, many proofs rely on *two-valued logic*, i.e., the maintained assumption that every proposition is either true or false. For example, the proof of a sufficiency theorem "if A, then B" often proceeds using *proof by contradiction*, as follows: Establish that the *falsity* of the proposition "if A, then B" would imply the *falsity* of a proposition C that is known (or assumed) to be true. Assuming two-valued logic, the proposition "if A, then B" must then be true, even if it is not possible to construct (calculate, generate, ...) the realization of B that corresponds to a realization of A.

In contrast, the sufficiency "examples" (if A, then B) established by a c-ABM modeler take the following externally constructive form: The c-ABM modeler conducts an experimental study with a c-ABM model that exhibits property A and observes that every resulting model outcome exhibits property B.

However, as stressed in Sect. 4.3, a c-ABM model can in fact be a *thickly constructive* blend of classical and constructive mathematics, in the following sense. Agent states consist of data, attributes, and/or methods. Agent attributes can include evolved or acquired non-constructive beliefs ("leaps of faith") as well as constructive beliefs based on input-output relationships directly observed or experienced in interactions with other agents. Agent methods can include belief-dependent behavioral rules for the determination of intended acts, i.e., acts the agents intend to express within their computational world. Agent interactions (expressed agent acts) depend on agent methods. Finally, all world events (changes in agent states) are driven by agent interactions. Consequently, world events can depend on a mix of non-constructive and constructive agent beliefs.

Does the thick constructivity of c-ABM necessarily imply that c-ABM is the "best" modeling approach for the study of real-world social systems whose human participants act on the basis of non-constructive as well as constructive beliefs? Absolutely not. However, as Fig. 4.1 depicts, human-subject experiments and c-ABM

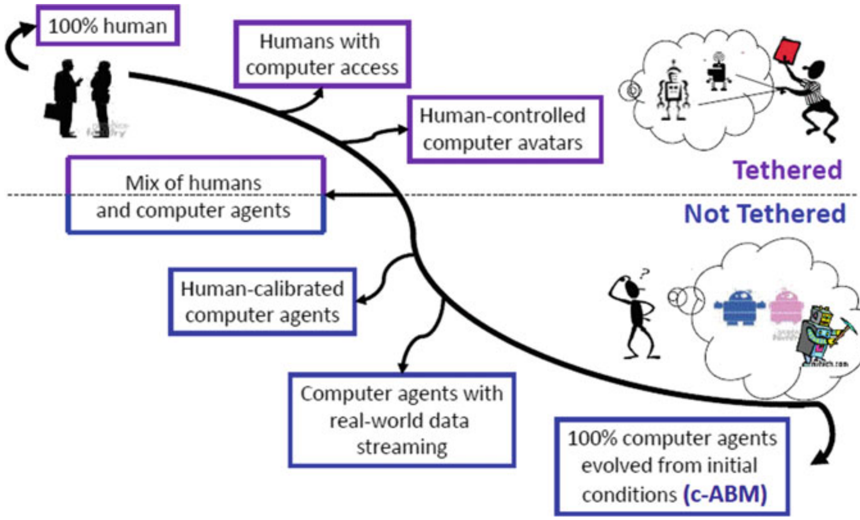


Fig. 4.1 A spectrum of experiment-based modeling methods ranging from 100% human subjects to 100% computer agents (c-ABM)

constitute the two polar end-points for a spectrum of hybrid human/agent experiment-based modeling methods that are well-suited for such studies.

The repository Tesfatsion (2022i) provides annotated links to studies focusing on possible synergies *between* human-based and agent-based experimental studies, as well as annotated links to several experiment-based studies (e.g., serious game research by Zheng and Gardner 2017) involving a *mix* of humans and computer agents. However, to date, the full range of hybrid human/agent experiment-based modeling methods depicted in Fig. 4.1 has not been systematically explored.

Acknowledgements I am very grateful to Shu-Heng Chen, Marcin Czupryna, and Volker Grimm for their extensive thoughtful comments on a previous version of this perspective. I am also thankful to the following people for helpful comments on related topics, many directed specifically to my coverage of these related topics in Tesfatsion (2021b): Richie Adelstein; W. Brian Arthur; Bob Axelrod; Rob Axtell; Costas Azariadis; Ross Baldick; Cesar Enrique Garcia Diaz; Marty Eichenbaum; Dan Friedman; Mauro Gallegati; Nick Gotts; Omar Guerrero; Dirk Helbing; Mike Honig; Dan Houser; Marco Janssen; George Judge; Lorenzo Kristov; Sergio Mariotti; Sheri Markose; Bob Marks; Vittorio Nespeca; Dawn Parker; Gary Polhill; Steve Railsback; Ryan Schoppe; Rajiv Sethi; Ken Steiglitz; Shyam Sunder; Peter Swann; Nick Vriend; Steve Widergren; and Mike Woodford.

References

Arthur WB (2021) Foundations of complexity economics. Nat Rev Phys 3:136–145. https://doi.org/10.1038/s42254-020-00273-3
Axelrod R (1984) The evolution of cooperation. Basic Books Inc, New York

Axtell RL, Farmer JD (2022) Agent-based modeling in economics and finance: past, present, and future, Journal of Economic Literature, to appear

Battula S, Tesfatsion L (2020) Agent-based modeling of electricity systems (AMES) - version 5.0. Github code/data repository, Battelle Memorial Institute. https://github.com/ames-market/AMES-V5.0

Chen S-H (2012) Varieties of agents in agent-based computational economics: a historical and an interdisciplinary perspective. J Econ Dyn Control 36(1):1–25

Chen S-H (2016) Agent-based computational economics: how the idea originated and where it is going. Routledge, New York

Chen S-H (ed) (2002) Evolutionary computation in economics and finance. Physica-Verlag, Heidelberg

Chen S-H, Li S-P (2012) Econophysics: bridges over a turbulent current. Int Rev Financ Anal 23:1–10

Chen S-H, Yeh C-H (2001) Evolving traders and the business school with genetic programming: a new architecture of the agent-based artificial stock market. J Econ Dyn Control 25(3–4):363–393

Dawid H, Delli Gatti D (2018) Agent-based macroeconomics, Chapter 2. In: Hommes C, LeBaron B (eds) Handbook of computational economics 4: heterogeneous agent modeling. Handbooks in economics series, North Holland (Elsevier), Amsterdam, pp 63–156

Dosi G, Roventini A (2019) More is different ... and complex! the case for agent-based macroeconomics. J Evolut Econ 29:1–37. https://doi.org/10.1007/s00191-019-00609-y

Epstein JM (2006) Generative social science: studies in agent-based computational modeling. Princeton University Press, Princeton

Epstein JM, Axtell RL (1996) Growing artificial societies: social science from the bottom up. The MIT Press, Cambridge

Gallegati M (2018) Complex agent-based models. New economic windows series, Springer International, Switzerland

Hofstadter DR (1983) Metamagical themas: computer tournaments of the Prisoner's Dilemma suggest how cooperation evolves. Sci Am 248(5):16–26

Judd KL (2006) Computationally intensive analyses in economics, chapter 17. In: Tesfatsion L, Judd KL (eds) Handbook of computational economics 2: agent-based computational economics. Handbooks in economics series, North Holland (Elsevier), Amsterdam, pp 881–893

Kalaba RE, Tesfatsion L (1991) Solving nonlinear equations by adaptive homotopy continuation. Appl Math Comput 41(2):Part II:99–115

Kirman A (2011) Complex economics: individual and collective rationality. The Graz Schumpeter lectures, Routledge, New York

LeBaron B, Tesfatsion L (2008) Modeling macroeconomies as open-ended dynamic systems of interacting agents. Am Econ Rev (Pap Proc) 98(2):246–250

Lori G, Porter J (2018) Agent-based modeling for financial markets. In: Chen S-H, Kaboudan M, Du Y-R (eds) The Oxford handbook of computational economics and finance. Oxford handbooks online. Oxford University Press, Oxford

Rahmandad H, Sterman J (2008) Heterogeneity and network structure in the dynamics of diffusion: comparing agent-based and differential equation models. Manag Sci 54(5):998–1014

Railsback SF, Grimm V (2019) Agent-based and individual-based modeling, 2nd edn. Princeton University Press, Princeton

Richiardi M (2013) The missing link: agent-based models and dynamic microsimulation, chapter 1. In: Leitner S, Wall F (eds) Artificial economics and self organization: agent-based approaches to economics and social systems, Springer Lecture Notes in Economics and Mathematical Systems 669, Berlin, Germany, pp 3–15

Rovelli C (2018) The order of time. Riverhead Books, New York, English translation by Simon Carnell and Erica Segre

Ryle G (1949) The concept of mind. Hutchinson, London

Sinitskaya E, Tesfatsion L (2015) Macroeconomies as constructively rational games. J Econ Dyn Control 61:152–182

Smolin L (2007) The trouble with physics: the rise of string theory, the fall of science, and what comes next. Houghton Mifflin, Boston
Tesfatsion L (2001a) Introduction to the special double issue on agent-based computational economics. J Econ Dyn Control 25(3–4):281–293
Tesfatsion L (2001b) Introduction to the special issue on agent-based computational economics. Comput Econ 18(1):1–8
Tesfatsion L (2001c) Guest editorial: agent-based modeling of evolutionary economic systems. IEEE Trans Evolut Comput 5(5):437–441
Tesfatsion L (2001d) Structure, behavior, and market power in an evolutionary labor market with adaptive search. J Econ Dyn Control 25(3–4):419–457
Tesfatsion L (2017) Modeling economic systems as locally-constructive sequential games. J Econ Methodoly 24(4):384–409
Tesfatsion L (2018) Electric power markets in transition: agent-based modeling tools for transactive energy support, chapter 13. In: Hommes C, LeBaron B (eds) Handbook of computational economics 4: heterogeneous agent models. Handbooks in economics series. North Holland (Elsevier), Amsterdam, pp 715–766
Tesfatsion L (2021a) A new swing-contract market design for wholesale power markets. IEEE press series on power engineering. Wiley Inc, Hoboken
Tesfatsion L (2021b) Agent-based modeling: the right mathematics for social science? Keynote address. In: 16th annual social simulation conference (virtual), Sponsored by the European social simulation association (ESSA), September 20–24. https://www2.econ.iastate.edu/tesfatsi/CompleteABM.SSC2021Keynote.LTesfatsion.pdf
Tesfatsion L (2022a) Agent-based computational economics (ACE): homepage. https://www2.econ.iastate.edu/tesfatsi/ace.htm
Tesfatsion L (2022b) ACE distributed news notes (1997–2017). https://www2.econ.iastate.edu/tesfatsi/acepast.htm
Tesfatsion L (2022c) ACE resource site: journal and publisher information. https://www2.econ.iastate.edu/tesfatsi/publish.htm
Tesfatsion L (2022d) Empirical validation and verification of agent-based models. https://www2.econ.iastate.edu/tesfatsi/empvalid.htm
Tesfatsion L (2022e) Learning and the embodied mind. https://www2.econ.iastate.edu/tesfatsi/aemind.htm
Tesfatsion L (2022f) Agent-based macroeconomics. https://www2.econ.iastate.edu/tesfatsi/amulmark.htm
Tesfatsion L (2022g) Agent-based financial economics. https://www2.econ.iastate.edu/tesfatsi/afinance.htm
Tesfatsion L (2022h) Agent-based electricity market research. https://www2.econ.iastate.edu/tesfatsi/aelect.htm
Tesfatsion L (2022i) Mixed experiments with real and computational agents. https://www2.econ.iastate.edu/tesfatsi/aexper.htm
Tesfatsion L (2022j) Agent-based computational economics: overview and brief history, economics working paper #21004, ISU digital repository (Originally posted 3/2021; Latest revision 5/2022), Iowa State University, Ames, IA 50011-1054. https://lib.dr.iastate.edu/econ_workingpapers/126
Tesfatsion L, Judd KL (eds) (2006) Handbook of computational economics 2: agent-based computational economics. Handbooks in economics series. North Holland (Elsevier), Amsterdam
Wilensky U, Rand W (2015) An introduction to agent-based modeling. MIT Press, Cambridge, MA, USA
Zheng RZ, Gardner MK (eds) (2017) Handbook of research on serious games for educational purposes. IGI Global, Hershey, PA, USA

Chapter 5
Sequential Monte Carlo Squared for Agent-Based Models

Thomas Lux

This Chapter is dedicated to Shu-Heng Chen on the occasion of his 60th birthday. Looking back on 25 years of academic exchange Shu has been a role model of a devoted scientist and a constant source of inspiration. The present celebration brings back lots of fond memories of various visits at Chengchi University and regular meetings at SCE and other conferences that always proved stimulating impulses to my own work. This chapter aims to provide a glance on how our joint topic of agent based modeling has developed from its theoretical beginning in the 90s to an empirically grounded field in current research.

Abstract This chapter proposes a combination of two sequential Monte Carlo algorithms for Bayesian estimation of agent-based models (ABM), one for the hidden state(s) of the system, and a second one for the parameters. The former consists of a meanwhile well-known particle approximation of the likelihood, while for the latter I implement a generic SMC algorithm designed along the lines of Del Moral et al. (2006). In an application to a well-studied financial ABM, the performance of this combined SMC approach compares well to frequentist estimation in terms of the precision of the mean of its posterior and its dispersion, and it offers a much less computationally demanding alternative to standard Markov chain Monte Carlo methods for Bayesian estimation.

Keywords Agent-based methods · Bayesian estimation · Sequential Monte Carlo

JEL Classification: G12 · C15 · C58

T. Lux (✉)
Department of Economics, University of Kiel, Olshausenstr. 40, 24118 Kiel, Germany
e-mail: lux@economics.uni-kiel.de

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_5

5.1 Introduction

With the increasing attractivity of agent-based modelling (ABM) in economics, validation and assessment of the goodness-of-fit of such models has become an important area of research (Chen et al. 2012). Over the years, practitioners of the ABM methodology have set out to explore a variety of alternative approaches to this end. One avenue that has first been pursued by many papers has been moment-based estimation. Given that Generalized Method of Moments (GMM) and Simulated Method of Moments (SMM) are also frequently used to validate econometric models with intractable likelihoods, this seemed an obvious choice within an economics framework. Examples of this strand of the literature include Franke (2009), Franke and Westerhoff (2011, 2012, 2016), Grazzini and Richiardi (2015), Ghonghadze and Lux (2016) and Chen and Lux (2018). A related broader approach has already been proposed much earlier for ABMs in ecology in the form of pattern-oriented modelling (Grimm et al. 2005). When patterns are understood as descriptive statistics, this approach would lead to moment-based estimates, but one can, of course, also imagine more qualitative information entering the validation of ABMs that would make an exact quantification cumbersome.

However, in applications in economics and finance, such qualitative criteria are usually absent and all empirical features one attempts to model are available in numerical form. The quest is then rather for the most efficient estimation (balancing precision of the parameter estimates and computational effort). Moment-based estimations typically score low on the efficiency scale. While one almost never has closed-form solutions for a likelihood function available in an ABM setting, there often remains the alternative of numerical approximations of the likelihood which should be bound to deliver more precise estimates than moment-based methods, albeit at the cost of higher use of computational resources. Such an approach has been recently pursued by Lux (2018) and Bertschinger and Mozzhorin (2021). Lux (2018) explores the use of a so-called particle filter, also known as an instant of a Sequential Monte Carlo (SMC) algorithm, both in a frequentist and in a Bayesian setting. In the particle filter approximation, the ABM is interpreted as a state-space model with some important properties of the agents (strategies, expectations, sentiment) constituting the hidden, unobservable variables. Inspection shows that economic ABMs can typically be cast into a general nonlinear state-space framework. In such a setting, the agent-specific details would contribute the latent part, whereas the measurements are the economic variables one aims to explain. Lux (2018) shows that this approach works well in a frequentist setting for two well-established simple ABMs of speculative interaction in financial markets.

When using the same approach for the numerical approximation of the likelihood in a Bayesian setting, the implementation within a Markov Chain Monte Carlo (MCMC) algorithm encounters problems: For one of these models, so low acceptance rates of the proposals in the Markov chain are obtained that its implementation becomes computationally unfeasible. For the other model, MCMC could be applied, but the computation time to obtain a Markov chain of reasonable size after a potentially long transitory phase was much higher than the resources needed for frequentist estimation. As a consequence, this paper only reported single MCMC runs, while in the frequentist sphere Monte Carlo simulations could be conducted to assess the finite-sample properties of the particle approximation. While theoretically, it is guaranteed that the Markov chain of a Particle MCMC (PMCMC) algorithm converges to the posterior under relatively general conditions (Andrieu et al. 2010), the computational demands are, nevertheless, troublesome. Lux (2022) has revisited this issue applying methods for adaptive choice of the proposal distribution within the PMCMC framework. A similar approach had already been pursued by Golightly and Wilkinson (2011) for ABMs in ecology. While adaptation alleviated the computational demands allowing to gear the acceptance rate towards its theoretical optimum, and allowing an MCMC implementation for all models considered, it also revealed certain principal limitations of the PMCMC approach: Namely, the particle approximation of the likelihood leads to stochastic fluctuations around the 'true' likelihood for each set of parameter values. If these fluctuations display relatively high variability (which depends on the model, its parameters and the underlying data), it will often generate realisations of the likelihood in the Markov chain that are not replaced by new draws for a long time. Hence, the chain tends to become degenerate. This problem can, in principle, be overcome by making the numerical approximation more precise which, however, requires an appropriate increase of the number of particles. Examples with empirical time series in Lux (2022) show how easily one encounters limits of computational feasibility when attempting to implement a PMCMC algorithm with sufficient mixing. Hence, while adaptation helps, it might also eventually hit hard boundaries of computational limitations.

This raises the question whether alternatives to standard MCMC exist that would make Bayesian estimation possible and easier to implement in the context of ABMs. This Chapter explores an Sequential Monte Carlo strategy designed for this purpose.

In general, SMC samples from a sequence of distributions. In a Bayesian setting it would start from the prior and attempt to approach the posterior by a number of steps through intermediate distributions. Since in all these steps, a particle filter (also subsumed under the SMC family of algorithms) is applied for each evaluation of the likelihood we can speak of an SMC-squared approach.

5.2 The Particle Filter Within an SMC Sampler

ABMs typically contain important latent, unobservable variables as well as variables that correspond to empirical measurements. With this distinction, an ABM can be categorized as a member of a very general class of state-space models.

Let us denote the parameter we wish to estimate by θ, with prior $p_0(\theta)$, the vector of latent variables at time t by x_t and the vector of observed variables by y_t and assume there is a latent process $f_\theta(x_{t+1}|x_t)$ and an observation process $g_\theta(y_t|x_t)$. Both processes might include stochastic components , but need not be represented by known conditional densities. Indeed, the only requirement that needs to be imposed is that the recursions $f(\cdot)$ and $g(\cdot)$ can be determined by some algorithm. Hence, any computer simulation that determines some relevant quantities from hidden interactions of artificial agents falls into this category. In many such models one would include observation noise in the process governing the observed variable, i.e. $g_\theta(y_t|x_t) = \tilde{g}_\theta(y_t|x_t) + \epsilon_t$, where $\tilde{g}(\cdot)$ provides the material link from the hidden variables to the observed ones, and ϵ_t is measurement noise unrelated to the economic interactions encapsulated in $\tilde{g}(\cdot)$ and $f(\cdot)$. For instance, in macroeconomics Dynamic Stochastic Equilibrium Models (DSGE) are fitted into the state-space format by the assumption that the key macro variables (GDP, consumption, etc.) are noisy observations of their 'true' hidden realisations (cf. Herbst and Schorfheide, 2015). If such innocuous measurement noise can be assumed, the structure of the model allows a numerical approximation of the likelihood of the data given a parameter set θ, irrespective of how complex the underlying mechanisms in $f(\cdot)$ and $\tilde{g}(\cdot)$ might be. We can then use the iterative form of the likelihood function:

$$L(y_1, ..., y_T|\theta) = P(y_1|\theta) \prod_{t=2}^{T} P(y_t|y_{t-1}, \theta),$$

with T the sample size of the observed data and $P(\cdot)$ the conditional or unconditional densities implied by the distribution of the measurement error, ϵ_t. The particle filter approximates the conditional densities in Eq. (5.1) by a 'swarm' of discrete samples:

$$P(y_t|y_{t-1}, \theta) \simeq \frac{1}{B} \sum_{j=1}^{B} P(y_t|x_t^{(j)})$$

with B the number of particles. The particle filter is applied as an SMC algorithm over the length of the time series conducting the following sequence of operations:

1. An initial swarm of particles $\{x_1^{(j)}\}$, $j = 1, \ldots, B$ is sampled from the stationary distribution of the latent variable x_t.
2. The likelihood for each particle is evaluated yielding $P(y_1|x_1^{(j)})$.
3. Given the relative likelihoods particles are resampled using multinomial draws with weights $w_1^{(j)} = \frac{P(y_1|x_1^{(j)})}{\sum_{l=1}^{B} p(y_1|x_1^{(l)})}$.
4. The new set of particles is iterated from $t = 1$ to $t = 2$ using the process $f(x_{t+1}|x_t)$.
5. The previous steps 2 through 4 are repeated for $t = 2$ to T.

This algorithm provides not only an approximation to the likelihood of the entire sample, but it also acts as a filter for the hidden variables x_t for which an estimate can be obtained for each time unit t at the end of step 2.

The above particle filter provides an estimate of the likelihood given a certain set of parameters, θ. When embedding this approximation in a frequentist estimation of the parameters, one has to resort to optimization algorithms for non-smooth objective functions because of the discreteness of the particles. In a Bayesian MCMC estimation, the estimated likelihood of the initial parameter set and that of the new proposal would enter into the usual expression for the probability of acceptance of the new draw. As has been mentioned above, this PMCMC algorithm converges to the same posterior as MCMC with exact likelihoods (Andrieu et al. 2010), but practical applications might impose unbearable computational costs to obtain a sufficiently mixing Markov chain.

As an alternative, one might explore to move from the prior to the posterior via a second SMC algorithm, now applied in the θ-dimension. To this end, we apply an example of a generic Sequential Monte Carlo sampler as formalized in Del Moral et al. (2006).

The second sampler for the parameters of the model works in the following way:

1. To initialize the algorithm at iteration $\tau = 1$ an initial swarm of N particles for the parameters, $\{\theta_1^{(k)}\}$, is sampled from the prior $p(\theta)$.
2. The likelihood $P(y_1, ..., y_T|\theta_1^{(k)}) = p(\mathbf{y_T}|\theta_1^{(k)})$ is determined for all particles using the particle filter algorithm described above.
3. Given the relative likelihoods particles are reweighted with weights $w_1^{(k)} = \frac{p(\mathbf{y_T}|\theta_1^{(k)})}{\sum_{n=1}^{N} p(\mathbf{y_T}|\theta_1^{(n)})}$.
4. A new set of particles is generated from the previous distribution using multinomial draws with weights $w_1^{(k)}$ and a transition kernel $k_1(\theta^*|\theta)$.
5. For iterations $\tau \geq 2$, step 2 is repeated for all particles $\{\theta_\tau^{(k)}\}$. Given the likelihood $P(\mathbf{y_T} \mid \theta_\tau^{(k)})$, particles are now resampled with weights:

$$w_t^{(k)} = \frac{P(\mathbf{y_T}|\theta_\tau^{(k)})}{\sum_{n=1}^{N} P(y_\tau|\theta_\tau^{(n)}) \sum_{l=1}^{N} w_{\tau-1}^{(l)} k_t(\theta_\tau^{(n)} \mid \theta_{\tau-1}^{(l)})}. \tag{5.1}$$

With these new probabilities, the former step 4 is implemented drawing a new set of particles using multinomial draws with these weights and a transition kernel $k_\tau(\theta^* \mid \theta)$.

The iterations are stopped once some termination criterion is fulfilled. The weights in Eq. 5.1 combine the usual relative likelihood with a term reflecting the genealogy of the particle from the previous set at iteration $\tau - 1$ via the transition kernel. The above algorithm is a generic SMC algorithm that constitutes an example of the general framework developed by Del Moral et al. (2006). Given its construction principle, it should therefore converge to the posterior of the parameters given the observations. While the first round already samples from the posterior, the subsequent ones can be seen as local explorations of the discrete distribution of round one. These serve to correct for Monte Carlo errors and are repeated as long as it takes to reach a stationary regime.[1]

[1] Our setting, thus, corresponds to what is described as a 'homogeneous sequence' in Del Moral et al. (2006). The sequential Monte Carlo framework also encompasses other designs such as a non-homogeneous procession towards the posterior as achieved by tempering to generate intermediate

For practical applications, a suitable stopping criterion has to be formulated. For this purpose, we conduct at every iteration a two-sample test comparing the joint distribution of the parameters and the likelihood between iteration $\tau - 1$ and τ. If the test does not reject the null hypothesis that both are samples from the same underlying distribution at a certain confidence level, we terminate the second SMC algorithm. The test adopted for this purpose is the two-sample test of Baringhaus and Franz (2004) that is based on a comparison of the Euclidean distance of the points of the combined sample with those of the two individual samples. A confidence level of 95% is used, and upon termination, the last two samples are combined to approximate the posterior distribution (since the null of them sharing the same underlying distribution could not be rejected).

If one would conceive as too high the computational demands of running as many iterations as needed for such convergence, one could impose a lower threshold for the required confidence level. It needs to be mentioned that the above approach is different from the SMC^2 algorithm proposed by Chopin et al. (2013) which has an interlaced development of the sampling of particles representing the x-dimension and those representing the θ-dimension. In contrast, the SMC^2 here consists of a particle filter used within every step of the second SMC approximation of the posterior distribution (indeed, the particle filter is applied here to each particle in the θ-dimensions generated in the course of the second SMC application). Preliminary experiments with the SMC^2 algorithm by Chopin et al. (2013) for the same test problem below did not show a clear advantage of one ore the other.

5.3 Test Example

Like in the previous applications of various SMC and MCMC algorithms to agent-based models (Lux 2018, 2022) we use the model of Alfarano et al. (2008) as our test bed for the above estimation. This model appears well-suited fo this purpose because (i) it is a genuine agent-based model meaning that we simulate its artificial time series via an ensemble of autonomous (but simple) agents whose behavior defines the latent process and influences the dynamics of the observation process (financial returns), (ii) this model satisfies properly the definition of a state-space process albeit with a complex ABM setting for the latent variable(s). In contrast, many other similar ABMs are not state-space processes proper, but can be rather characterized as belonging to the more general class of dynamic models with latent states as they entail a feedback effect from the observations to the latent variables.

The latent part of the model consists of an ensemble of N_c investors prone to changes between bullish and bearish sentiment. With sentiment being conceived as a binary variable, their changes of attitude are formalized by transition rates:

---

distributions between the prior and the posterior (cf. Ogundijo and Wang 2017). Comparison of the efficiency of these alternatives is left for future research.

$$\pi_{+,t} = a + bn_{+,t},$$
$$\pi_{-,t} = a + bn_{-,t}$$

with $n_{+,t}$ ($n_{-,t}$) the number of current members of the bullish (bearish) group, $n_{+,t} + n_{-,t} = N_c$, and the first transition rate capturing changes from "$-$" to "$+$", and vice versa for the second.

The overall sentiment at time t is defined as $x_t = \frac{n_{+,t} - n_{-,t}}{N_c}$ just as one defines this measure in survey expectations among market participants. Asset returns r_t are driven by both fundamentals and sentiment according to the following simple law of motion:

$$r_t = \epsilon_t + x_t - x_{t-1}$$

with $\epsilon_t \sim N(0, \sigma_f^2)$ the fundamental news. Any further parameters for the weights of both factor of influence have been suppressed since the three parameters a, b and σ_f provide already for enough flexibility to capture many scenarios of different relative importance of fundamentals and sentiment.

5.4 Monte Carlo Simulations

We now provide an illustration of the performance of the above SMC-squared approach to Bayesian estimation of agent-based models. Our aim is to estimate the joint posterior distribution of the set of parameters $\theta = \{a, b, \sigma_f\}$. The length of our underlying time series is T = 2,000 integer observations extracted from continuous discrete event simulations of the model presented in Sect. 5.3. We use both B = 1,000 and N = 1,000 particles for the states and the parameters, respectively, and also compare this benchmark with simulations using only 520 particles for either the states or the parameters, or both. The computational demands of this method depend on the number of iterations required to activate the stopping criterion. As it turns out, the time to complete the simulations are still quite a bit higher than under a frequentist maximum likelihood estimation based on the particle filter (the first part of the present algorithm) together with a Nelder-Mead optimization. However, on the positive side, the computational demands are much lower than under a MCMC approach (Lux 2018, 2022) even though the latter has used adaptive adjustments of the transition kernel. One major advantage of the SMC-squared estimation is that it can be easily parallelized. In the results reported in Table 5.1, parallelization on forty cores has been used for the second SMC step for the parameter particles.

Underlying parameters of the Monte Carlo simulations summarized in Table 5.1 are: a $=$ 0.0003, b $=$ 0.0014, $\sigma_f = 0.03$. Uniform priors have been used for all parameters with support [0, 0.005] for parameters a and b and [0, 0.5] for σ_f. The transition kernel of the second SMC sampler for the parameters has been specified as a random walk kernel with covariance matrix equal to two times the covariance matrix of the previous sample of particles. Given the still relatively heavy computational

Table 5.1 Statistics of posterior distribution

B = 520, N = 520				B = 520, N = 1000		
	Mean	stdc	Median	Mean	stdc	Median
a	0.884	0.557	0.639	0.674	0.300	0.599
b	1.639	0.317	1.571	1.683	0.322	1.623
σ_f	31.278	2.104	31.039	30.632	1.610	30.542
tsec	2,152	2,119	1,495	6,598	4,128	5,951
iter.	9.85	11.51	6.50	14.85	10.11	13.00
B = 1000, N = 520				B = 1000, N = 1000		
	Mean	stdc	Median	Mean	stdc	Median
a	0.904	0.558	0.696	0.642	0.290	0.475
b	1.580	0.236	1.622	1.555	0.174	1.562
σ_f	29.112	1.831	29.415	29.826	1.171	30.326
tsec	5,080	3,662	3,978	17,546	13,990	12,144
iter.	10.80	9.16	8.00	19.60	15.48	14.40

Table 5.1: The table shows the mean estimates over 20 Monte Carlo runs of the SMC^2 algorithm, i.e., the mean over the 20 runs of the mean of the posterior together with the standard error of these means across the 20 Monte Carlo runs, as well as the mean of the 20 medians of the same runs. Also shown are the statistics of the mean computation time in seconds (tsec) and the average number of iterations. The numbers of B and N have been chosen as multiples of 40 (the numbers of cores used in the parallel implementation of the algorithm). The parameters have been multiplied by 10^3 for better readability

demands, only 20 simulations have been conducted. When comparing the results for the mean parameter estimates and their dispersion, we find that they compare well with the frequentist results reported in Lux (2018) for the same numbers T and B (N, of course, has no counterpart as it is replaced by optimization via the Nelder-Mead algorithm). The comparison shows that b is estimated with higher precision, which comes at the cost of a lower precision for parameter a. For the third parameter σ_f, the present algorithm scores only slightly better. The results of Table 5.1 also highlight that a better approximation of the likelihood (higher B) is more important than a higher number of θ-particles, although increasing the latter also has a positive effect in our limited set of Monte Carlo simulations. The time needed for the Nelder-Mead optimization is only about 500 s per run with a standard deviation of only 50. In contrast, the present algorithm despite parallelization needs almost 35 times as long for convergence, and the dispersion of the computing time across different runs is much higher. Indeed, the number of iterations for convergence has been varying between 6 and 56.

Figure 5.1 shows one typical example at the higher end (convergence at iteration 50). Displayed are the means and the inner 95% quantiles as they develop over the iterations for the parameters and the likelihood. For the likelihood, the first two rounds have been skipped as they differ by orders of magnitudes from those observed from round 3 onward. There is one particular feature which is shared by all simulations: From the very broad prior the distribution shrinks to a very narrow one concentrated

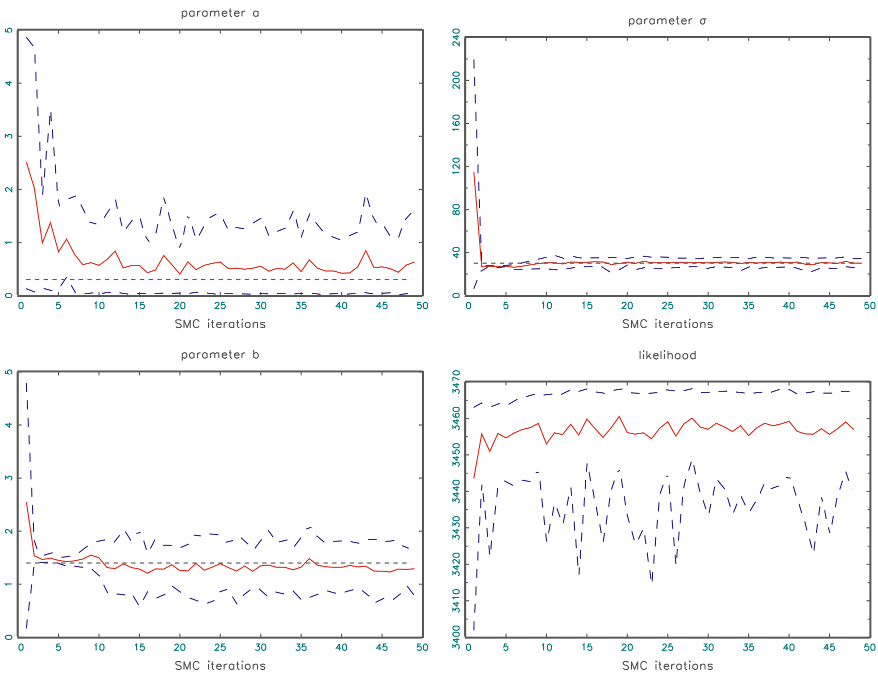


Fig. 5.1 Evolution of the mean and inner 95% quantile of the particles of the parameters and of the likelihood in one exemplary run of the SMC-squared algorithm. The broken lines indicate the 'true' values of the parameters. Parameters have been multiplied by 10^3 for better readability

on the vicinity of the 'true' values for b and σ_f. This typically happens within the first 4 to 5 rounds. Subsequently, the distribution becomes broader again, and exhibits very little changes of its shape, so that at some point between iteration 6 to 10, for the naked eye the variation appears less systematic. This also applies to parameter a although is does not display this non-monotonic behavior of the dispersion of the distribution. While it takes 50 periods in the present example until the two-sample test does not reject identity of the distribution of two subsequent samples any more, one might approximate the posterior already at an earlier step without too much loss of precision. Figure 5.2 provides a somewhat more complete image of the evolving distributions exhibiting their inter-quantile ranges and full extent from the minimum to the maximum values. The inspection shows that rejections before round 50 might mostly be due to fluctuations in the parameters a and b that show wide variation of their inter-quartile ranges.

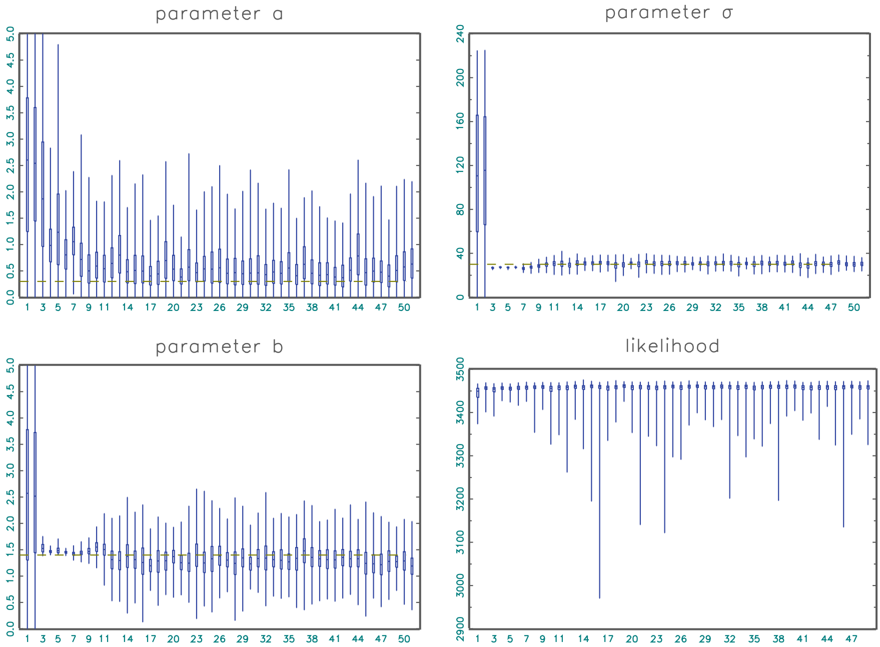


Fig. 5.2 Evolution of the distribution of the particles of the parameters and of the likelihood for the same exemplary run as depicted in Fig. 5.1. The boxplots show the inter-quartile range and the full width of the distributions

5.5 Conclusion

This Chapter has proposed an alternative to standard MCMC for Bayesian estimation of agent-based models that have found to be extremely computation intensive and sometimes even unfeasible given usual computational resources, cf. Lux (2018, 2022). The proposed SMC^2 estimation combines an SMC sampler for the hidden states with another one working on the particles.

In contrast to the well-known SMC^2 algorithm of Chopin et al. (2013), the two parts are, however, not executed simultaneously, but the first one provides an approximation to the likelihood function as an input for the second SMC sampler. A comparison between these different design principles would be an obvious task for future research.

A small Monte Carlo exercise has shown that the means of the posteriors, and their dispersion compare well with the point estimates of a frequentist estimation with the same number of particles for the states. While the SMC^2 approach is still more computation-intensive than frequentist estimation, it, of course, comes with the added benefit of a complete characterization os the posterior distribution of the parameters.

References

Alfarano S, Lux T, Wagner F (2008) Time variation of higher moments in a financial market with heterogeneous agents: an analytical approach. J Econ Dyn Control 32(1):101–136
Andrieu C, Doucet A, Holenstein R (2010) Particle Markov Chain Monte Carlo methods. J R Stat Soc: Ser B (Stat Methodol) 72(3):269–342
Baringhaus L, Franz C (2004) On a new multivariate two-sample test. J Multivar Anal 88(1):190–206
Bertschinger N, Mozzhorin I (2021) Bayesian estimation and likelihood-based comparison of agent-based volatility models. J Econ Interac Coord 16(1):173–210
Chen S-H, Chang C-L, Du Y-R (2012) Agent-based economic models and econometrics. Knowl Eng Rev 27(2):187–219
Chen Z, Lux T (2018) Estimation of sentiment effects in financial markets: a simulated method of moments approach. Comput Econ 52:711–744
Chopin N, Jacob PE, Papaspiliopoulos O (2013) SMC^2: an efficient algorithm for sequential analysis of state space models. J R Stat Soc: B 75(3):397–426
Del Moral P, Doucet A, Jasra A (2006) Sequential Monte Carlo samplers. J R Stat Soc: B 68(3):411–436
Franke R (2009) Applying the method of simulated moments to estimate a small agent-based asset pricing model. J Empir Financ 16(5):804–815
Franke R, Westerhoff F (2011) Estimation of a structural stochastic volatility model of asset pricing. Comput Econ 38(1):53–83
Franke R, Westerhoff F (2012) Structural stochastic volatility in asset pricing dynamics: estimation and model contest. J Econ Dyn Control 36(8):1193–1211
Franke R, Westerhoff F (2016) Why a simple herding model may generate the stylized facts of daily returns: explanation and estimation. J Econ Interac Coord 11(1):1–34
Ghonghadze J, Lux T (2016) Bringing an elementary agent-based model to the data: estimation via GMM and an application to forecasting of asset price volatility. J Empir Financ 37:1–19
Golightly A, Wilkinson DJ (2011) Bayesian parameter inference for stochastic biochemical network models using particle Markov Chain Monte Carlo. Interface Focus 1(6):807–820
Grazzini J, Richiardi M (2015) Estimation of ergodic agent-based models by simulated minimum distance. J Econ Dyn Control 51:148–165
Grimm V, Revilla E, Berger U, Jeltsch F, Mooij WM, Railsback SF, Thulke H-H, Weiner J, Wiegand T, DeAngelis DL (2005) Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science 310(5750):987–991
Herbst EP, Schorfheide F (2015) Bayesian estimation of DSGE models. Princeton University Press
Lux T (2018) Estimation of agent-based models using sequential Monte Carlo methods. J Econ Dyn Control 91:391–408
Lux T (2022) Bayesian estimation of agent-based models via adaptive Particle Markov Chain Monte Carlo, Working paper, Comput Econ 60:451–477
Ogundijo OE, Wang X (2017) Bayesian estimation of scaled mutation rate under the coalescent: a sequential Monte Carlo approach. BMC Bioinf 18(1):1–15

Chapter 6
Toward a General Model of Financial Markets

Nihad Aliyev and Xue-Zhong He

Abstract This paper discusses the idea of reconciling efficient market hypothesis and behavioral finance using the literature of decision theories and information sciences. The focus is centered on the precision and reliability of information and the broad definition of rationality. The main thesis advanced is that the roots of behavioral anomalies are the imprecision and reliability of information. We propose a general framework that subsumes efficient and inefficient markets as special cases. We also show that the proposed framework helps to understand behavioral anomalies.

Keywords Efficient markets · Behavioral finance · Decision theory · Information uncertainty

JEL Classification G02 · G10 · G14 · D81

Is our expectation of rain, when we start out for a walk, always more likely than not, or less likely than not, or as likely as not? I am prepared to argue that on some occasions none of these alternatives hold, and that it will be an arbitrary matter to

---

This chapter is based on/draws from Nihad Aliyev's 2019 Ph.D. thesis, titled 'Financial Markets with Multidimensional Uncertainty', submitted to UTS Business School, University of Technology, Sydney. We thank Rafik Aliev, David Easley, Doug Foster and Lotfi. A. Zadeh for their comments and the participants of conferences at Quantitative Methods in Finance, Australasian Finance & Banking Conference and seminar at University of Technology Sydney. The financial support of Australian Research Council (DP130103210) and UTS Quantitative Finance Research Center is gratefully acknowledged.

N. Aliyev (✉)
University of Technology Sydney, Finance Department, UTS, Building 8, 14/28 Ultimo Rd, Ultimo 2007, Australia
e-mail: nihad.aliyev@uts.edu.au

X.-Z. He
International Business School Suzhou (IBSS) Business Building (BS), South Campus, Xi'an Jiaotong-Liverpool University, No. 8 Chongwen Road, Suzhou, People's Republic of China
e-mail: xuezhong.he@xjtlu.edu.cn

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_6

decide for or against the umbrella. If the barometer is high, but the clouds are black, it is not always rational that one should prevail over the other in our minds, or even that we should balance them, though it will be rational to allow caprice to determine us and to waste no time on the debate.

Keynes (1921) "A Treatise on Probability" [p. 31].

6.1 Introduction

Financial market participants face multiple dimensions of uncertainty about the fundamental values of securities and other market characteristics. To explain asset price movements without an obvious cause (e.g., external news), Romer (1993) models investors' uncertainty about the quality of other investors' information. More recently, Banerjee and Green (2015) show that uncertainty about the informativeness of other investors' signals leads to a nonlinear price that reacts asymmetrically to news. Similarly, Aliyev et al. (2022) show that modeling uncertainty about the composition of traders and quality of traders' information along with the fundamental values of securities generates a rich characterization of the dynamics of security prices in response to order flow and provides intuition about the prevalence of liquidity and price crashes. A common approach in modeling multi-dimensional uncertainty in these models is that uncertainty is quantified probabilistically and market participants make trading decisions based on expected utility theory of Von Neumann and Morgenstern (1944).

Departing from the fully probabilistic characterization of uncertainty, Dow and Werlang (1992) model optimal investment decisions with expected utility under a non-additive probability measure (e.g., Schmeidler 1989), distinguishing between quantifiable "risks" and unknown "uncertainties".[1] Easley et al. (2013) investigate the effect of ambiguity about hedge fund investment strategies on asset prices and aggregate welfare. More recently, Aliyev and He (2022b) model price formation with ambiguous liquidity provision to explain sudden liquidity changes in financial markets and Aliyev and He (2022a) model ambiguity about the composition of traders.

While these models provide useful characterizations of different dimensions of uncertainty in financial markets, they require some structure to obtain closed-form or numerical solutions. In this paper, we depart from the mainstream economic modeling approach and discuss the complexity of the real-world information and the

[1] The idea of unknown uncertainty (ambiguity) dates back to Knight (1921) and Keynes (1921), where they distinguish between risk (when relative odds of the events are known) and uncertainty (when the degree of knowledge only allows the decision maker to work with estimates). Ellsberg (1961) provides experimental evidence to the ideas of Knight and Keynes. The behavior of ambiguity-aversion has been popularized in the decision making context by Choquet expected utility of Schmeidler (1989) and maxmin expected utility of Gilboa and Schmeidler (1989). Since then different approaches have been taken to model ambiguity (e.g., Bewley 2002; Klibanoff et al. 2005).

implications for financial markets. We discuss information in the broadest possible way that lends itself to possible quantitative scrutiny. Specifically, we use Zadeh (2011) classification—numerical, interval-valued, second-order uncertain, fuzzy and Z information—based on the reliability and precision of information.

We argue that individuals are subjectively rational if they apply correct decision technique to each class of information separately rather than defining rationality based on only one decision technique such as the standard expected utility theory. The general framework that we propose contains multiple decision theories corresponding to each class of information. Using two of the decision theories in our framework, we rationalize buying insurance (choosing certainty in preference to uncertainty) and lottery tickets (choosing uncertainty in preference to certainty) simultaneously and equity premium puzzle (difference between the average equity and bond returns).

Consequently, our thesis is that efficient market hypothesis and behavioral finance become special cases of this framework with the specificity and reliability of information connecting them. When information is precise and reliable, we expect everyone to act rationally, whereas when information becomes imprecise and unreliable, we expect the opinions to diverge and create a room for irrationality. Put differently, when information is imprecise and unreliable, the environment becomes too complex to be explained by the standard techniques and we refer behaviors that escape the standard techniques as irrationality.

A general view of market efficiency that we present in this paper is different from the one presented by efficient market hypothesis and behavioral finance and subsumes both efficient and inefficient markets as its special cases. The beliefs represented by conceptually different theories at different information classes separately support efficient market hypothesis and behavioral finance. Thus, in our framework, efficient and inefficient markets have a certain degree of truthness depending on the prevalence of different information classes in financial markets.

6.1.1 Efficient Market Hypothesis

Efficient market hypothesis proposes that large price movements result from the arrival of new fundamental information into the market. This is first comprehensively formalized in Osborne (1959) by making a number of assumptions. One of the underlying assumptions of Osborne's world is the "logical decision", meaning investors are assumed to form expectations probabilistically and choose the course of action with a higher expected value. That is to say, investors form objective probabilities and make rational decisions as if they know each individual outcome.

Another crucial assumption made by Osborne (1959) is an "independence of decisions" in the sequence of transactions of a single stock which leads to independent, identically distributed successive price changes. This implies that changes in prices can only come from unexpected new information. In this setting, central limit theorem assures that price changes converge to a Gaussian distribution. The Gaussian distribution assumption is later generalized by Mandelbrot (1963) to stable Paretian distribution to account for the empirical evidence of leptokurtic distributions of price changes.

Fama (1965) provides further empirical support for the stable Paretian distribution hypothesis. Fama (1965) also argues that an independence assumption may still hold due to the existence of sophisticated traders (the so-called smart money), even though the processes generating noise and new information are dependent. An independence assumption is therefore consistent with efficient markets where prices at every point in time represent the best estimates of intrinsic values. The combination of independence and stable Paretian distribution allows Fama to argue that the actual prices adjust instantaneously to the changes in intrinsic value due to the discontinuous nature of the stable Paretian distribution. This version of efficient market hypothesis includes random walk theory as a special case.

The first formal general economic argument of efficient markets is, however, given by Samuelson (1965) by focusing on the martingale property established by Bachelier (1900). Similar to the "logical decision" assumption of Osborne (1959), Samuelson (1965) also assumes that people in financial markets make full use of the past probability distribution.

In summary, the proponents of this paradigm essentially argue the use of a probability calculus as a foundation of the economic analysis. This view also justified an application of probability based decision techniques to financial markets and became associated with the modern portfolio theory of Markowitz (1952a) and capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965) and Mossin (1966).

6.1.2 Behavioral Finance

The starting point of behavioral finance stands in sharp contradiction to the logical decisions assumption of efficient market hypothesis in forming expectations. The early works of Kahneman and Tversky is a foundational block of this area of finance. In a series of experiments, they show that people use heuristic to decide under uncertainty and conjecture that the same heuristic plays an important role in the evaluation of uncertainty in real life. In their seminal paper, Kahneman and Tversky (1979) present a critique of the expected utility theory and develop the prospect theory as an alternative.

Inspired by Kahneman and Tversky's works, Shiller (1979) shows that long-term interest rates are too volatile to be justified by rational models. Similarly, Thaler (1980) argues that consumers do not follow economic theory and proposes an alternative descriptive theory on the basis of the prospect theory. Shiller (1981) also argues that stock prices fluctuate too much to be justified by subsequent dividend changes.

All of these findings sharply contradicted the efficient market hypothesis and shaped the emergence of a new field. Finally, Bondt and Thaler (1985) marked the birth of behavioral finance with empirical evidence of overreaction hypothesis suggested by the experimental psychology. Since then, the number and magnitude of anomalies noticed by researchers have increased (although some are the mere results of data dredging).

The main arguments of behavioral finance are categorized by Shefrin (2000) as follows:

- Financial practitioners commit errors due to relying on rules of thumb.
- Frame of a decision problem influences financial practitioners' decisions.
- Heuristic-driven biases and framing effects lead the prices in financial markets to deviate from fundamental values.

Overall, efficient market supporters criticize behavioral finance for not having any unifying principles to explain the origin of behavioral anomalies and behavioral finance supporters criticize efficient market hypothesis for making unrealistic assumptions and systematic errors in predicting human behavior. In this paper, we examine the possibility of a general framework where both paradigms coexist. An exposition of a more general view of financial markets is the main purpose of this paper.

The rest of the paper is organized as follows. Section 6.2 presents a general framework to define subjective rationality as a broad concept. Section 6.3 briefly reviews candidate decision theories to account for the subjectively rational behavior. In Sect. 6.4, we discuss "insurance and gambling" and "equity premium" puzzles. Section 6.5 presents a novel representation of market efficiency. Section 6.6 concludes. Appendices include additional mathematical details of decision theories.

6.2 Broad Concept of Subjective Rationality

The main goal of this section is to answer a question of what is meant by *"rational"* behavior when the decision maker (DM) is confronted with different types of information. The proposed framework is based on the premise that the correct decision method changes when the specificity and reliability of information change.

We consider a subjective notion of rationality in the sense that the decision maker cannot be convinced that he is wrong. Gilboa, Maccheroni, Marinacci and Schmeidler (2010) show how the Knightian decision theory of Bewley (2002) and the maxmin expected utility (MEU) of Gilboa and Schmeidler (1989) are complementary to each other in terms of defining objective and subjective rationality. They argue that a choice is objectively rational if the DM can convince others that he is right in making the decision and subjectively rational if others cannot convince the DM that he is wrong in making the decision.

The current interpretation of rationality in economics and finance relies heavily on the subjective expected utility (SEU) axioms of Savage (1954). Simply, the DM is rational if he follows axioms of SEU and irrational if he does not. This definition of rationality is too narrow to capture a real life decision situation. In addition, this definition contradicts what has been argued by many economists such as Keynes (1921), Knight (1921), Shackle (1949), Arrow (1951) to name a few.

Specifically, Knight (1921) makes a clear distinction between risk (when relative odds of the events are known) and uncertainty (when the degree of knowledge only allows the DM to work with estimates). Arrow (1951) notes that descriptions of uncertain consequences can be classified into two major categories, those which use exclusively the language of probability distributions and those which call for some other principle. There is a need of another principle to be a supplement to a language of probability to better approximate a real life decision situation. The information that decisions are based on is not only uncertain in nature, but at the same time imprecise and partially true. Using only a probabilistic approach is not sufficient to treat uncertainty, imprecision and partial truthness of information adequately.

One of the main arguments advanced in information science literature is that imprecision of the real life information is possibilistic rather than probabilistic in nature and a fuzzy set theory is a necessary mathematical tool to deal with the possibilistic uncertainty (e.g., Zadeh 1978). While there has been substantial progress in modeling uncertainty probabilistically, economics as a discipline has been somewhat reluctant to account for the latter over the years.

By adopting the concept of subjective rationality in Gilboa et al. 2010, we consider a decision theoretic approach to discuss the right decision methods with different types of information. By classifying the problems according to the information types of a DM we find candidate decision theories to account for the subjective rationality. In that sense, irrationality in our context is a mismatch of the decision methodology and the information environment. For example, an application of an objective probability-based decision method to the problem with objective probability distribution is rational, while the same method is not avaliable in the situation where a DM has imprecise information. Thus, the expected utility theory (EUT) of Von Neumann and Morgenstern (1944) can be regarded as a right decision method in probability theory applicable circumstances as it builds upon objective probabilities.

By generous stretch of imagination, we use a similar logic to determine the right decision methods for more general classes of information. The higher the generality of information and corresponding decision theories, the more financial phenomena we can explain that we otherwise label as paradoxes (anomalies). In our framework we follow Zadeh (2011) who outlines the following classification of information based on its generality.

- **Numerical Information (Ground Level—'G').** This is single valued information with exact probability, e.g., there is 80% chance that there will be 3% growth in Australian economy next year.
- **Interval-valued Information (First Level—'F').** This is the first order uncertainty in which probability and value take intervals, e.g., there is 75–85% chance that there will be 2–3.5% growth in Australian economy next year.
- **Information with second-order uncertainty (Second Level—'S').** This is partially reliable information with sharp boundaries, e.g., there is 70–85% chance that there will be 1.5–3.5% growth in Australian economy next year and the lower probability of the given chances being reliable is 80%.

- **Fuzzy Information (Third Level—'T').** This is the information with un-sharp boundaries, e.g., there will be a moderate growth in Australian economy next year.
- **Z-information and visual information (Z Level—'Z').** This is partially reliable information with un-sharp boundaries and often in natural language, e.g., it seems likely that there will be a moderate growth in Australian economy next year.

The distinction between these levels can be difficult. Sometimes we can think of one upper level as the same as one below. However, there is no doubt that the degree of informativeness or specificity of information diminishes as we move away from the ground level. The former implies the latter while it is not true for the reverse.[2] We do not address when the information is sufficiently informative to serve a particular purpose. However, we approximate the definition of rationality in Table 6.1. Individuals can be considered subjectively rational along the diagonal in this framework. That is, for each level of information class there is a decision theory (methodology) to account for the subjectively rational behavior. Following the representation in Table 6.1, we define subjective rationality as follows.

Definition 1 A subjectively rational decision is consistent with different decision theories for different classes of information.

The representation of subjectively rational behavior in Table 6.1 hides a philosophical subtlety in itself. Philosophically, rationality is not a 0/1 property. One can therefore modify Table 6.1 to describe a degree of irrationality of the DM. For example, an irrationality of applying decision theory 1 to interval-valued information and applying the same method to Z information should be different—the latter is more irrational than the former. One can consistently apply the same logic to the upper right triangle of Table 6.1.

Note that, this is different from maximally rational, rational and minimally rational classification of Rubinstein (2001). What we have in mind is to set the maximum level as rational in the underlying information class and then reduce by one unit for each level of deviation from the underlying information class to describe the degree of irrationality. Also, we do not confine ourselves to only maxmin expected utility model in defining subjective rationality as proposed by Gilboa et al. 2010. This in turn enables us to differentiate irrationality of the decision maker. A changing degree of irrationality can provide a novel foundation on the theory of choice under uncertainty.

Similarly, an application of the more general decision theory where the less general is sufficient to capture the given decision situation is an inefficient decision making. One can consistently apply this logic to the lower left triangle of Table 6.1. In terms of consistency of the framework, an application of the more general decision theory should give the same result as the less general one in the corresponding information class of the less general theory. Nevertheless, the latter provides computational ease and saves the decision maker from an inefficient decision making.

---

[2] One can think of this generalization as a subsethood relation, $G \subseteq F \subseteq S \subseteq T \subseteq Z$, but not in a strict mathematical sense of subsethood since, for example, it is not obvious to see the relation between S and T.

Table 6.1 Subjective rationality This table defines subjective rationality as a correct method corresponding to each class of information. In this framework, there are five information classes and corresponding five decision theories to account for the subjectively rational behavior. The diagonals in the table correspond to subjectively rational decision-making. The upper part of the diagonal is irrational decision making and the lower part of the diagonal is inefficient decision making

	Information classes				
Method	Numerical information	Interval-valued information	Second-order uncertainty	Fuzzy information	Z information
Decision theory 1	Rational	–	–	–	–
Decision theory 2	–	Rational	–	–	–
Decision theory 3	–	–	Rational	–	–
Decision theory 4	–		–	Rational	–
Decision theory 5	–	–	–	–	Rational

In this framework, there are two fundamental reasons to move from one decision theory to another. First, a more general decision theory is needed if it solves at least one more paradox that the existing theory cannot solve. Second, the existing decision theory becomes inconvenient or excessively complex at some stage and it is desirable to move to a more convenient theory. The principle of replacing the existing decision theory with a more general decision theory is similar to the principle of requisite generalization in generalized information theory (GIT) (e.g., Klir 2005). A generalization is not optional, but requisite, imposed by the nature of the decision situation.

6.3 Candidate Decision Theories

In this section, we discuss candidates for the correct decision theories corresponding to five classes of information. To illustrate the differences of decision theories at different information classes, we consider a DM who chooses among three alternatives (bonds—f_1, stocks—f_2 and term deposit—f_3) for a short-term investment plan.

Table 6.2 Utilities of each act under different states of the economy This table shows the decision maker's precise utilities of three alternatives (bonds—f_1, stocks—f_2 and term deposit—f_3) under four states of economy strong growth (s_1), moderate growth (s_2), stable economy (s_3) and recession (s_4)

	States of the economy			
Acts	Strong growth (s_1)	Moderate growth (s_2)	Stable (s_3)	Recession (s_4)
Bonds (f_1)	15	9	8	4
Stocks (f_2)	16	9	4	0
Term deposit (f_3)	10	10	10	10

6.3.1 Decision Theory 1

Suppose, the DM evaluates each alternative under strong growth (s_1), moderate growth (s_2), stable economy (s_3) and recession (s_4). Table 6.2 shows the precise utilities under each state of the economy for different acts.

The DM also has perfect information about the uncertainty of the states with the following (subjective) probabilities: $P(s_1) = 0.5$, $P(s_2) = 0.3$, $P(s_3) = 0.15$, and hence, $P(s_4) = 0.05$. Clearly, the information of the DM is numerical and he is in the province of probability theory. For this type of simplistic information, classical measure and integral are adequate tools to calculate expected utilities of each act. In this environment, the DM can easily determine his preferences as $f_1 \succ f_2 \succ f_3$ by calculating expected utilities as

$$U(f_1) = 11.6, \quad U(f_2) = 11.3 \quad \text{and} \quad U(f_3) = 10. \tag{6.1}$$

In expected utility theory (EUT) of Von Neumann and Morgenstern (1944) and subjective expected utility (SEU) of Savage (1954), choice under uncertainty is perceived as the maximization of the mathematical expectation of individual utilities with respect to (subjective) probabilities. If preferences of the DM coincide with what is suggested by EUT or SEU, then his action is perfectly justifiable and can be regarded as rational based on the proposed framework. The optimal choice for the DM is f_1 (bonds). In what follows, we illustrate information classes where the classical tools are not directly applicable.

6.3.2 Decision Theory 2

Suppose, the DM still evaluates each alternative under strong growth (s_1), moderate growth (s_2), stable economy (s_3) and recession (s_4) and he notes the same utilities shown in Table 6.2. However, he now assigns the following subjective probability intervals: $P(s_1) = [0.4, 0.45]$, $P(s_2) = [0.3, 0.35]$, $P(s_3) = [0.15, 0.20]$, and hence, $P(s_4) = [0, 0.15]$.

The information of the DM now corresponds to the interval-valued information in Table 6.1.[3] Given the set of states S and its power set $\mathcal{F}(S)$, let $I = \{[l(s_i), u(s_i)] \mid i \in \mathbb{N}_4\}$ denote 4-tuples of probability intervals on $s_i \in S$, where $l(s_i)$ and $u(s_i)$ denote the lower and upper probability bounds of state i. Let $\mathcal{M}$ denote a convex set of probability distribution functions p

$$\mathcal{M} = \{p \mid l(s_i) \leq p(s_i) \leq u(s_i), i \in \mathbb{N}_4, \sum_{s_i \in S} p(s_i) = 1\}. \tag{6.2}$$

From the probability distributions in set $\mathcal{M}$, the lower probability measure (lower prevision) is defined for all $A \in \mathcal{F}(S)$ as $\eta(A) = \inf_{p\in\mathcal{M}} \sum_{x_i \in A} p(x_i)$. It follows from this definition that lower probabilities satisfy the conditions of capacities (i.e., monotone measures). Then, the lower probability measure η is

$$\eta(A) = \max\left\{\sum_{x_i \in A} l(x_i), 1 - \sum_{x_i \notin A} u(x_i)\right\}, \forall A \in \mathcal{F}(S). \tag{6.3}$$

For clarity, let us exemplify,

$$\eta\big(\{s_1, s_4\}\big) = \max\{l(s_1) + l(s_4), 1 - u(s_2) - u(s_3)\} = 0.45, \tag{6.4}$$

$$\eta\big(\{s_1, s_3, s_4\}\big) = \max\{l(s_1) + l(s_3) + l(s_4), 1 - u(s_2)\} = 0.65, \tag{6.5}$$

and report the values of lower probability measure for all states in Table 6.3.[4]

Based on the probability intervals, the DM can determine his preferences as $f_1 \succ f_3 \succ f_2$ by first ordering utility values in a descending order and then aggregating utilities with the Choquet integral with respect to the lower probability measure η.[5]

For a given alternative, the Choquet (expected) utility is computed as

$$\begin{aligned} U(f_i) =& \Big(u\big((f_i)(s_1)\big) - u\big((f_i)(s_2)\big)\Big)\eta(\{s_1\}) \\ &+ \Big(u\big((f_i)(s_2)\big) - u\big((f_i)(s_3)\big)\Big)\eta(\{s_1, s_2\}) \end{aligned}$$

---

[3] For simplicity we assume that only his probability assessments take interval values. This can be extended to interval-valued utilities in the sense of Gul and Pesendorfer (2014).

[4] Note that, $\mathcal{M}$ is non-empty if and only if, $\sum_{i=1}^{4} l(s_i) \leq 1$ and $\sum_{i=1}^{4} u(s_i) \geq 1$. Also, Eq. (6.3) is only applicable when I satisfies $\sum_{j\neq i} l(s_j) + u(s_i) \leq 1$ and $\sum_{j\neq i} u(s_j) + l(s_i) \geq 1$. These conditions are trivially satisfied when $l(s_4) = 1 - \sum_{i=1}^{3} u(s_i)$ and $u(s_4) = 1 - \sum_{i=1}^{3} l(s_i)$.

[5] Note that the use of lower probability measure is implicitly justified with an ambiguity-aversion.

Table 6.3 Lower probability measures for all states This table shows the lower probability measures of all states in the economy. We compute the lower probability measure following Eq. (6.3) using interval probabilities of the DM as $P(s_1) = [0.4, 0.45]$, $P(s_2) = [0.3, 0.35]$, $P(s_3) = [0.15, 0.20]$ and $P(s_4) = [0, 0.15]$

	States				
Lower probability	$\{s_1\}$	$\{s_2\}$	$\{s_3\}$	$\{s_4\}$	$\{s_1, s_2\}$
$\eta(A)$	0.4	0.3	0.15	0	0.7
	$\{s_1, s_3\}$	$\{s_1, s_4\}$	$\{s_2, s_3\}$	$\{s_2, s_4\}$	$\{s_3, s_4\}$
$\eta(A)$	0.55	0.45	0.45	0.35	0.20
	$\{s_1, s_2, s_3\}$	$\{s_1, s_2, s_4\}$	$\{s_1, s_3, s_4\}$	$\{s_2, s_3, s_4\}$	$\{S\}$
$\eta(A)$	0.85	0.80	0.65	0.55	1

$$+\Big(\big(u(f_i)(s_3)\big) - u\big((f_i)(s_4)\big)\Big)\eta(\{s_1, s_2, s_3\})$$
$$+\, u\big((f_i)(s_4)\big)\eta(S), \tag{6.6}$$

given that $u\big((f_i)(s_1)\big) \geq u\big((f_i)(s_2)\big) \geq u\big((f_i)(s_3)\big) \geq u\big((f_i)(s_4)\big)$.[6] Following the same steps for f_1, f_2 and f_3, we obtain the expected utilities of each alternative as

$$U(f_1) = 10.5, \quad U(f_2) = 9.7 \quad \text{and} \quad U(f_3) = 10, \tag{6.7}$$

leading to the preference order of $f_1 \succ f_3 \succ f_2$.

Due to the imprecise nature of probability intervals, the classical measure and integral become deficient to directly determine the optimal choice. Therefore, we first determine the convex set of probability distribution functions from the given intervals and calculate the lower envelope of this closed convex set as a lower probability measure. As this lower probability measure satisfies the conditions of Choquet capacities (or non-additive probabilites), the Choquet integral becomes the right tool to determine the expected utilities of each act. This is the Choquet Expected Utility (CEU) proposed by Schmeidler (1989) using the notion of non-additive probabilities.[7]

---

[6] The Choquet expectation is equivalent to adding the probability gap to the belief about the worst case scenario.

[7] With the convex capacities it is also well-known that CEU coincides with MEU under the assumption of ambiguity-aversion (see Proposition 3 of Schmeidler (1986) for proof). A capacity η is convex for all events $A, B \in \mathcal{F}(S)$ if it satisfies $\eta(A \cup B) + \eta(A \cap B) \geq \eta(A) + \eta(B)$.

6.3.3 Decision Theory 3

Suppose the DM evaluates each alternative under $S = \{s_1, s_2, s_3, s_4\}$ with the same precise utilities shown in Table 6.2. He also assigns the same interval probabilities: $P(s_1) = [0.4, 0.45]$, $P(s_2) = [0.3, 0.35]$, $P(s_3) = [0.15, 0.20]$ and $P(s_4) = [0, 0.15]$. This time a probability interval of $[0.7, 0.8]$ is assigned to measure an imprecise degree of confidence of the assigned probabilities. The traditional methods are also incapable of solving this problem due to the probability intervals and the second-order uncertainty imposed by the reliability of the assigned probabilities. There are two approaches to determine the optimal choice with second-order uncertain probabilities.

Approach 1. A direct way to address the problem is to use the methodology of interval-valued information and use the reliability of the information as the accuracy of the choice. With this approach, the order of preference is as in decision theory 2 ($f_1 \succ f_3 \succ f_2$) with an accuracy of $[0.7, 0.8]$.

To see the merit of this approach, consider the following illustration by Shafer (1987). Suppose, we have asked Fred if the streets outside are slippery. He replies "Yes" and we know that 80% of the time he speaks truthfully and 20% of the time he speaks carelessly, saying whatever comes into his mind. With

$$p_1 = \text{“the streets are slippery”}, \tag{6.8}$$

$$p_2 = \text{“the streets are not slippery”} \tag{6.9}$$

propositions, Shafer derives a belief of 0.8 in proposition $\{p_1\}$ and 0.2 in $\{p_1, p_2\}$.

If we don't have additional information, we should not allocate the remaining 0.2 between p_1 and p_2. In our example, the Shafer's illustration suggests that there are $[0.7, 0.8]$ units of evidence supporting $f_1 \succ f_3 \succ f_2$, and $[0.2, 0.3]$ units of evidence supporting all other combinations of preference order. In line with Shafer's example, with this approach, the utility of each alternative is

$$U(f_1) = 10.5, \quad U(f_2) = 9.7 \quad \text{and} \quad U(f_3) = 10 \tag{6.10}$$

with the reliability (confidence) of $[0.7, 0.8]$.

Approach 2. It also seems that the reliability of information and the ambiguity attitude of the DM are related to each other. It seems reasonable to assume that there is an inverse relationship between reliability and ambiguity-aversion. As the information gets more and more unreliable the DM becomes more ambiguity averse. In that sense, ambiguity-aversion α is a function of the reliability of information $\alpha = \psi(\underline{r}, \bar{r})$, where $\underline{r}$ and $\bar{r}$ denote the lower and upper reliability of information. We cannot determine what confidence function (ψ) would account for rational behavior. This is similar to the problem of which utility function makes the most sense. In that sense, any confidence function leading to ambiguity-aversion would suffice for our purposes. Suppose,

$$\psi(r) = 1 - \left(\frac{\underline{r} + \bar{r}}{2}\right). \tag{6.11}$$

Then, ambiguity-aversion, α, equals 0.25 and an application of α-MEU in the sense of Ghirardato et al. (2004)

$$U(f_i) = \alpha \min_{P \in \mathcal{M}} \int_S u(f_i(S))dP + (1-\alpha) \max_{P \in \mathcal{M}} \int_S u(f_i(S))dP, \tag{6.12}$$

results in

$$U(f_1) = 11.25, \quad U(f_2) = 10.79 \quad \text{and} \quad U(f_3) = 10, \tag{6.13}$$

which leads to the preference order of $f_1 \succ f_2 \succ f_3$.[8]

6.3.4 Decision Theory 4

Now consider the DM who notes the following trends under $S = \{\tilde{s}_1, \tilde{s}_2, \tilde{s}_3, \tilde{s}_4\}$.

(i) $\tilde{f}_1$ will yield high income under $\tilde{s}_1$, medium income under $\tilde{s}_2$, less than medium income under $\tilde{s}_3$ and small income under $\tilde{s}_4$;
(ii) $\tilde{f}_2$ will yield very high income under $\tilde{s}_1$, medium income under $\tilde{s}_2$, small income under $\tilde{s}_3$ and a notable loss under $\tilde{s}_4$;
(iii) $\tilde{f}_3$ will yield approximately the same medium income in all 4 fuzzy states of economy.

The possible set of states for the risk attitude of DM is $H = \{h_1, h_2\}$, where h_1 and h_2 (non-fuzzy in this example) stands for risk-aversion and risk-seeking, respectively. The DM also has information that $\tilde{s}_1$ will take place with a medium probability, $\tilde{s}_2$ will take place with a less than medium probability, $\tilde{s}_3$ with a small probability and $\tilde{s}_4$ with a very small probability.[9] The probability of the DM's risk-aversion is known to be about 70% and he is assumed to be risk-seeking when he is not risk-averse.

Approach 1. The first approach to make decision with fuzzy information is to compute a fuzzy-number-valued lower prevision and use the Choquet integral with respect to the computed lower prevision to calculate the total utility values of each act. This is essentially a generalized version of Choquet expected utility (CEU) of Schmeidler (1989) and the argument advanced by Aliev et al. 2012. Appendix A provides the details of formal mathematical treatment of fuzzy information.

---

[8] Instead, one can also consider the smooth decision making model of Klibanoff, Marinacci and Mukerji (2005),

$$U(f_i) = E_\mu[\phi(E_\pi(u(\cdot)))], \tag{6.14}$$

when an objective probability measure π and its subjective relevance μ are precise. However, if the uncertainties are imprecise in both, first and second order, one is left to use some imagination to solve similar problems.

[9] We use the notation of superimposed tilde for fuzzy values.

Approach 2. The second approach is more behavioral in nature. Because of capturing interaction among behavioral determinants to account for the fundamental level dependence of human behavior, behavioral decision-making with combined states under imperfect information (BDM) of Aliev et al. (2013) is another candidate to determine the optimal choice with this information class. BDM captures an interaction among factors induced by the fuzzy environment. In Appendix B, we apply BDM to decision making under fuzzy information and obtain the preference order as $\tilde{f}_2 \succ \tilde{f}_1 \succ \tilde{f}_3$.

6.3.5 *Decision Theory 5*

Now suppose the DM evaluates the same alternatives under the same economic conditions $S = \{\tilde{s}_1, \tilde{s}_2, \tilde{s}_3, \tilde{s}_4\}$ and his information is the same as the one outlined in decision theory 4. This time, he has a degree of reliability (expressed in a natural language) of the given information. Specifically, he is *very sure* about the outcomes of his three actions and he is *sure* about his probability assessment.

At this information class, the DM has imprecise and at the same time partially true information (Z-information). Following the arguments set forth in decision theory 3, there are also two approaches of solving the optimal solution with Z infomation.

Approach 1. The first approach is to use BDM and interpret the reliability of the given information as the accuracy of the preference order. The argument of decision theory 3 with the illustration of Shafer (1987) applies with the same logic. With this approach, the resulting preference order is $\tilde{f}_2 \succ \tilde{f}_1 \succ \tilde{f}_3/sure$.

Approach 2: The second approach uses the concept of Z-number suggested by Zadeh (2011). Formally, a Z-number is defined as an ordered pair $\hat{Z} = (\tilde{A}, \tilde{B})$ of fuzzy numbers to describe a value of a variable. Here, $\tilde{A}$ is an imprecise constraint on values of the variable and $\tilde{B}$ is an imprecise estimation of reliability of $\tilde{A}$. We refer to Aliev et al. (2015) for the arithmetic of Z-numbers and to Aliev et al. 2016 for the general theory of decisions (GTD). The GTD uses the idea of combined states argument of BDM and develops a unified decision model which subsumes most of the well-known decision theories as its special cases.

6.4 Paradoxes and Rationality

In modern economics literature, there is a lot of evidence contradicting the preference of Savage's axioms as well as the theory itself as a valid representation of rationality. The evidence ranges from Ellsberg (1961) to Kahneman and Tversky (1979). However, over the years, the compiled evidence is regarded as an irrationality of economic agents while Savage's axioms retained its normative ground in economics and finance. Hence, different paradigms such as efficient market hypothesis and behavioral finance are formed. An approach on the basis of imprecision and reliability of information enables us to judge when a subjectively rational belief obeys the probability calculus and when it is less structured.

This type of rationality is not the first time introduced to the economics and finance literature. For example, the response of Thaler (1980) to Friedman and Savage (1948) expert billiard player analogy essentially suggests that acting on the basis of prospect theory may sometimes be judged as rational. The flopping of a fish analogy of Lo (2004) suggests that the same motion (flopping) makes a fish rational in one environment (underwater) and makes it irrational in another environment (dry land). Our analysis reveals the information-based picture of these environments. We next revise two of the existing well-known paradoxes of behavioral finance.

6.4.1 Insurance and Gambling

Buying both insurance and lottery tickets is a norm rather than an exception and it is hard to reconcile with classical rational decision making. Buying insurance means a DM chooses a certainty in preference to uncertainty, whereas buying a lottery ticket suggests choosing uncertainty in preference to certainty. To see how insurance and gambling can be rationalized by BDM, consider the following example.

Alice considers to buy fire insurance for her house. She notes *a small loss* (insurance premium) under $\tilde{s}_1$ (no fire) and *a very large gain* under $\tilde{s}_2$ (fire) if she buys the insurance ($\tilde{f}_1$). She also notes *a very large loss* under $\tilde{s}_2$, while *nothing* happens under $\tilde{s}_1$ if she does not buy the insurance ($\tilde{f}_2$). Fire occurrs with a *very small* probability, $\tilde{P}(\tilde{s}_2)$. Table 6.4 summarizes the gains and losses of Alice under different states.

Alice also considers to buy a lottery ticket in the hope of winning the mega jackpot. As illustrated in Table 6.5, she notes *a very small loss* (ticket price) under $\tilde{s}'_1$ (not win) and *a very large gain* under $\tilde{s}'_2$ (win) if she buys a ticket ($\tilde{f}'_1$). She feels *nothing* if she does not buy a ticket ($\tilde{f}'_2$). Probability of winning $\tilde{P}(\tilde{s}'_2)$ is *very small*. The probability of her risk-aversion $\tilde{P}(h_1)$ is 70% and she is known to be risk-seeking (h_2) when she is not risk-averse.

Friedman and Savage (1948) suggest an S-shaped utility function to rationalize this behavior and the approach is criticized by Markowitz (1952b). The value function

Table 6.4 Gains or losses under different states of buying and not buying fire insurance This table shows the gains or losses of buying ($\tilde{f}_1$) and not buying ($\tilde{f}_2$) fire insurance under different states (no fire—$\tilde{s}_1$ and fire—$\tilde{s}_2$)

	No fire—$\tilde{s}_1$	Fire—$\tilde{s}_2$
Buy—$\tilde{f}_1$	Small loss	Very large gain
Not buy—$\tilde{f}_2$	No loss/gain	Very large loss

Table 6.5 Gains or losses under different states of buying and not buying lottery ticket This table shows the gains or losses of buying ($\tilde{f}_1$) and not buying ($\tilde{f}_2$) a lottery ticket under different states (no win—$\tilde{s}'_1$ and win—$\tilde{s}'_2$)

	$\tilde{s}'_1$ (no win)	$\tilde{s}'_2$ (win)
$\tilde{f}'_1$ (buy)	A very small loss	A very large gain
$\tilde{f}'_2$ (don't buy)	No loss/gain	No loss/gain

in BDM also follows from the Prospect Theory. Therefore, one can follow the steps of BDM outlined in Appendix B and verify that Alice should not be ashamed of buying both a lottery ticket and fire insurance at the same time.

6.4.2 Equity Premium Puzzle

The equity premium puzzle (first noted by Mehra and Prescott (1985)) refers to the large difference between the average equity returns and average returns of a fixed interest bearing bonds. To see how equity premium puzzle can be rationalized consider the following example.

Bob is an (only) investor with an initial wealth of W_0 who can invest in two assets, a risky stock with a price of p and an uncertain payoff R_s in state $s \in S$, and a bond with a unit price and a certain payoff R. Suppose he invests a units of stock and b units of bond. Further denote $\pi(s)$ as an additive probability distribution over the state s. The end-of-period wealth is $W_s = R_s \cdot a + R \cdot b$. Using a budget constraint, $W_0 = p \cdot a + b$, we obtain the end-of-period wealth as $W_s = R \cdot W_0 + [R_s - R \cdot p] \cdot a$. First consider Bob as expected utility theory (EUT) maximizer as a benchmark case,

$$U(W_1, ..., W_s) = \sum_{s \in S} \pi_s \cdot u(W_s) = \sum_{s \in S} \pi_s \cdot u(R \cdot W_0 + [R_s - R \cdot p] \cdot a). \quad (6.15)$$

Without loss of generality, for an equilibrium stock price of p_0 with a total investment in stock, $a > 0$, and bonds, $b = 0$, Bob maximizes his total utility

$$U'(a) = \sum_{s \in S} \pi_s \cdot [R_s - R \cdot p_0] \cdot u'(R_s \cdot a) = 0. \quad (6.16)$$

The stock price p_0 follows from the first order condition in Eq. (6.16) as

$$p_0 = \frac{\sum_{s \in S} \pi_s \cdot R_s \cdot u'(R_s \cdot a)}{R \cdot \sum_{s \in S} \pi_s \cdot u'(R_s \cdot a)}. \quad (6.17)$$

The benchmark equity premium is thus given by

$$\tau(p_0) = \frac{\sum_{s\in S} \pi(s) \cdot R_s}{p_0 \cdot R}. \tag{6.18}$$

Now suppose, preferences of Bob are represented by α-MEU in the sense of Ghirardato et al. (2004) as

$$U(W_1, ..., W_s) = \alpha \cdot \min\{u(W_1), ..., u(W_s)\} + (1 - \alpha) \cdot \max\{u(W_1, ..., u(W_s)\}, \tag{6.19}$$

where α denotes Bob's ambiguity-aversion. Denote $\bar{R} = \max\{R_1, ..., R_s\}$ and $\underline{R} = \min\{R_1, ..., R_s\}$. Then, for $a > 0$, the total utility of investing in a amount of stock is

$$U(a) = \alpha \cdot u(R \cdot W_0 + (\underline{R} - R \cdot p) \cdot a) + (1 - \alpha) \cdot u(R \cdot W_0 + (\bar{R} - R \cdot p) \cdot a). \tag{6.20}$$

Similarly, without loss of generality, for an equilibrium stock price of p with a total investment in stock, $a > 0$, and bonds, $b = 0$, Bob maximizes his total utility as

$$U'(a) = \alpha \cdot u'(\underline{R} \cdot a) \cdot (\underline{R} - R \cdot p) + (1 - \alpha) \cdot u'(\bar{R} \cdot a) \cdot (\bar{R} - R \cdot p) = 0. \tag{6.21}$$

The equilibrium stock price p follows from the first order condition in Eq. (6.21) as

$$p = \frac{\alpha \cdot u'(\underline{R} \cdot a)\underline{R} + (1 - \alpha) \cdot u'(\bar{R} \cdot a)\bar{R}}{\alpha \cdot R[u'(\underline{R} \cdot a) - u'(\bar{R} \cdot a)] + u'(\bar{R} \cdot a) \cdot R}. \tag{6.22}$$

The equity premium when Bob has $\alpha-$MEU preferences is given by

$$\tau(p) = \frac{\sum_{s\in S} \pi(s) \cdot R_s}{p \cdot R}. \tag{6.23}$$

To compare the equity premium in the benchmark EUT case in Eq. (6.18) and the $\alpha-$MEU case in Eq. (6.23), we first assume Bob is risk neutral (i.e., $u'(\cdot) = c$).

- **Case 1 (a).** The equity premium increases with the ambiguity aversion of the investor.

It follows from Eqs. (6.18) and (6.23) that it is sufficient to show that the equilibrium stock price decreases with the ambiguity-aversion α due to an inverse relationship between the equity premium and the stock price. For $u'(\cdot) = c$, the equilibrium stock price in Eq. (6.22) reduces to

$$p = \frac{\alpha \cdot (\underline{R} - \bar{R}) + \bar{R}}{R}. \tag{6.24}$$

Since $\underline{R} < \bar{R}$, the equilibrium stock price will be lower and the equity premium will be higher the more ambiguity averse the investor is.

- **Case 1 (b).** For $\alpha > 1/2$, the equity premium $\tau(p)$ exceeds the benchmark $\tau(p_0)$ when the expected return exceeds the average of the minimum and maximum returns, i.e.,

$$\frac{(\underline{R} + \bar{R})}{2} < E_\pi[R_s]. \tag{6.25}$$

For $u'(\cdot) = c$, the benchmark stock price in Eq. (6.17) reduces to

$$p_0 = \frac{\sum_{s \in S} \pi_s \cdot R_s}{R} = \frac{E_\pi[R_s]}{R}. \tag{6.26}$$

Combining Eqs. (6.24) and (6.26) obtains

$$p = p_0 + \frac{1}{R} \cdot \big[\alpha \cdot \underline{R} + (1 - \alpha) \cdot \bar{R} - E_\pi[R_s]\big]. \tag{6.27}$$

The necessary and sufficient condition for $p < p_0$ and $\tau(p) > \tau(p_0)$ follows from Eq. (6.27) as

$$\alpha \cdot \underline{R} + (1 - \alpha) \cdot \bar{R} < E_\pi[R_s], \tag{6.28}$$

which reduces to

$$\frac{(\underline{R} + \bar{R})}{2} < E_\pi[R_s] \tag{6.29}$$

for $\alpha = 1/2$. Since $\tau(p)$ increases in α (Case 1(a)), the condition must be true for $\alpha > 1/2$.

- **Case 1 (c).** For $\alpha = 1$ (full ambiguity-aversion), the equity premium with ambiguity $\tau(p)$ always exceeds the benchmark $\tau(p_0)$

  It follows from Eq. (6.24) that for $\alpha = 1$,

$$p = \frac{\underline{R}}{R} < \frac{E_\pi[R_s]}{R} = p_0, \tag{6.30}$$

leading to $\tau(p) > \tau(p_0)$.

We now consider the second case in which Bob is risk-averse with a strictly decreasing marginal utility function. Then, the following result is immediate.

- **Case 2.** The equity premium $\tau(p)$ of a risk-averse (i.e., $u'(\cdot) < 0$) and ambiguity-averse (i.e., $\alpha = 1$) investor exceeds the benchmark equity premium $\tau(p_0)$ of a risk-averse investor.

In this scenario, the benchmark stock price is given by Eq. (6.17). For the risk- and ambiguity-averse investor with $\alpha = 1$, however, the stock price is given by

$$p = \frac{\underline{R}}{R} \tag{6.31}$$

Comparing Eqs. (6.17) and (6.31) obtains $p_0 > p$, and therefore, $\tau(p) > \tau(p_0)$ following

$$\frac{\sum_{s\in S}\pi_s\cdot R_s\cdot u'(R_s\cdot a)}{\sum_{s\in S}\pi_s\cdot u'(R_s\cdot a)} > \underline{R}. \tag{6.32}$$

With this example, we have only added an ambiguity attitude of the DM in the sense of Ghirardato et al. (2004) to his risk attitude and confirmed that ambiguity-aversion requires an incremental equity premium. In this context, any pessimistic attitude adds an incremental requirement for the equity premium and completes the pieces of the puzzle.[10] So deeply rooted is our commitment to EUT and SEU, that we regard such patterns as paradoxical or irrational. As we move away from the ground level of the information many of the paradoxes, it turns out, can be rationalized by more general decision theories.

6.5 Fuzzy Representation of Market Efficiency

The main difference between efficient market hypothesis and behavioral finance comes down to the difference between probability (classical set theory) and possibility (fuzzy set theory). These theories are at the two extreme sides of Table 6.1. Probability calculus lies in the foundation of efficient markets, whereas fuzzy set theory lies in the foundation of behavioral finance.[11] Since fuzzy sets are the generalized version of classical sets, the concept of market efficiency is a fuzzy concept. In this section, we represent the market efficiency as a fuzzy concept.

In essence, our argument thus far is that the room for behavioral finance increases as we move from left to right in Table 6.1. We summarise our arguments as follows.

(i) Imprecision and reliability of information lead to different opinions.
(ii) Imprecision and reliability of information lead to “behavioral biases” in financial decision-making.
(iii) Imprecision and reliability of aggregate information in financial markets lead to more “behavioral anomalies” observed in the market.

The first point is an economic primitive. The second point is a natural extension of the first point. The third point is the economy-wide aggregate of the second point. Note that our points are different from the argument of Friedman (1953) that in aggregate noise traders cancel each other. Here we do not conjecture the survival of noise traders in financial markets. We rather focus on the imprecise and unreliable information that is received by everyone and perceived differently.

[10] The results of this simple example with a representative investor carry over to more general models of financial markets (e.g., Epstein and Schneider 2008; Ju and Miao 2012). These results are also consistent with liquidity dry-ups due to ambiguity premium on the bid-ask spread in Aliyev and He (2022b) and composition uncertainty premium in Aliyev and He (2022a) when traders are sufficiently uncertainty averse.

[11] Peters (1996) discusses how fuzzy membership functions can be used to understand some of the prominent behavioral biases.

These points are intuitive. When market participants receive fully-reliable simple numerical information (pure fact), then it seems reasonable to assume that there is homogeneous belief in the market. However, when the received information is fuzzy (e.g., medium growth) and partially true (e.g., sure), different perceptions of a natural language lead to different opinions, "behavioral biases", and in aggregate, "behavioral anomalies". In our framework with a concept of subjective rationality, we are not in favor of calling them as "biases" or "anomalies". We argue that when information becomes imprecise and unreliable, behavioral factors become one's only strength to play with (as noted by Keynes at the beginning quote of this paper). They are called biases and anomalies because it is extremely odd to explain them with the probabilistic calculus.

These points also reveal a general view of market efficiency that is different from what is presented by the two main paradigms of finance. The beliefs represented by conceptually different theories at the two edges of the information classes separately support efficient market hypothesis and behavioral finance. Therefore, it seems to us efficient and inefficient markets supported by the fundamentally different beliefs have a certain degree of truthness depending on the dominance of different information classes in financial markets.

Figure 6.1 graphically illustrates the degree of market efficiency corresponding to different price levels. In the figure, when asset prices vary between P_1 and P_4, efficient and inefficient (undervalued or overvalued) markets become special cases

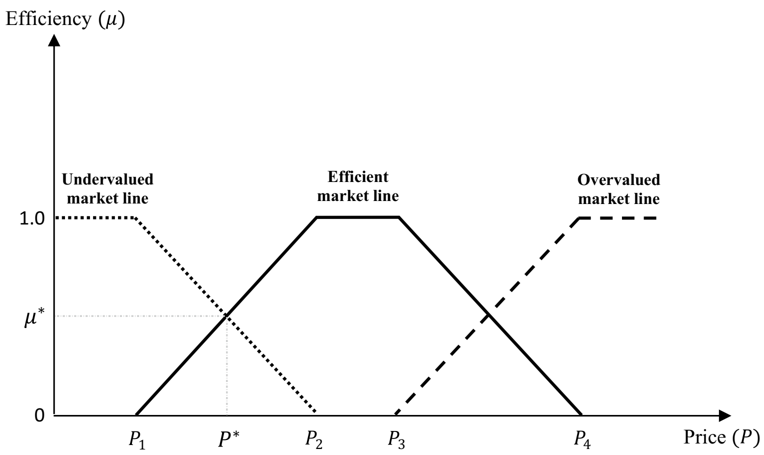


Fig. 6.1 Fuzzy market efficiency This figure illustrates the degree of market efficiency. The dotted line is an undervalued market line, the solid line is an efficient market line and dashed line is an overvalued market line. When price P is less than P_1 the degree of undervaluation is 1; when $P_1 < P < P_2$, the degree of undervaluation and the degree of efficiency are between 0 and 1. When $P_2 < P < P_3$, the degree of efficiency is 1. When $P_3 < P < P_4$, the degree of efficiency and the degree of overvaluation are between 0 and 1. When $P > P_4$, the degree of overvaluation is 1

of this framework without sharp boundaries. Specifically, Figure 6.1 subsumes 3 specific fuzzy sets—undervalued market, efficient market and overvalued market—with the degree of membership to each set in the y axis and the price level on the x axis. When the asset price is less than P_1 the degree of undervaluation (or membership to the undervaluation set) is 1; when $P_1 < P < P_2$, the degree of undervaluation and the degree of efficiency are between 0 and 1. When $P_2 < P < P_3$, the degree of efficiency is 1 and so on.

There are two subtle reasons that this direction to be distinguished from both efficient markets and behavioral finance paradigms. First, the existence of efficient markets is not entirely excluded in this framework as in Grossman and Stiglitz (1980). This is also different from the current weak-form, semi-strong form and strong-form market efficiency concepts. Second, behavioral finance regards the probability based decision techniques as its superior, though the subjective rationality in our argument can take behavioral factors into account when the information that the decision is based on becomes imprecise and partially reliable.

6.6 Conclusion

The exposition of a more general view of financial markets in this paper shows how decision theory and information science literature can be used to better understand financial markets. We conclude with the following three remarks.

(i) Classical probability theory favors efficient market hypothesis and rules out a possible room for behavioral finance.
(ii) Individual behavioral "biases" and aggregate market "anomalies" mainly originate from the imprecision and reliability of information.
(iii) Decision theories built on the foundations of capacities, bi-capacities, set of probabilities, and in a more general setting, fuzzy set theory can be used to complement the probability based decision theories to reconcile efficient market hypothesis and behavioral finance.

Appendix A—Mathematical Preliminaries

Capacities and Choquet Integral. Capacities replace the additivity requirement of classical measures with a less restrictive requirement of monotonicity. Let Ω be a universal set characterizing the *states of nature* and $\mathcal{F}(\Omega)$ its non-empty power set with appropriate algebraic structure characterizing the *events*. A capacity is a real-valued set function, $\eta(A)$, defined on the set A of events $\mathcal{F}(\Omega)$ that is normalized ($\eta(\emptyset) = 0$, $\eta(\Omega) = 1$) and monotonic (for all A, B in $\mathcal{F}(\Omega)$, $A \subseteq B \Rightarrow \eta(A) \leq \eta(B)$). Additional continuity conditions (below and above) are required when Ω is infinite. Suppose $\Omega = \{\omega_i\}_{i=1}^{n+1}$ is finite. Without loss of generality we can rank a non-negative (utility) function $f(\omega_k)$ on Ω as $f(\omega_1) \geq f(\omega_2) \geq ... \geq f(\omega_n)$ and $f(\omega_{n+1}) = 0$. Then, the expected value (i.e. Choquet integral) of f on Ω with respect to a capacity η is expressed as

$$E_\eta[f] = \sum_{k=1}^{n} (f(\omega_k) - f(\omega_{k+1}))\eta(\{\omega_1, \omega_2, ..., \omega_k\}). \tag{A.1}$$

The mathematical treatment of Choquet capacities can be found in Choquet (1955), Dempster (1967), Shafer (1976), and Schmeidler (1986, 1989).

Bi-Capacities and Choquet-like Aggregation. Bi-capacities are natural generalization of capacities in the context of decision-making where underlying scales are bipolar as in the prospect theory. Let $\mathcal{Q}(\Omega) := \{(A, B) \in \mathcal{F}(\Omega) \times \mathcal{F}(\Omega) \mid A \cap B = \emptyset\}$ denote the set of all pairs of disjoint sets. A bi-capacity is a real-valued set function, $\eta(A, B)$, defined on $\mathcal{Q}(\Omega)$ that is normalized ($\eta(\emptyset, \emptyset) = 0$, $\eta(\Omega, \emptyset) = 1 = -\eta(\emptyset, \Omega)$) and monotonic (for all A, B in $\mathcal{Q}(\Omega)$, $A \subseteq B \Rightarrow \eta(A, \cdot) \leq \eta(B, \cdot)$ and $\eta(\cdot, A) \geq \eta(\cdot, B)$).[12] Suppose $\Omega = \{\omega_i\}_{i=1}^{n+1}$ is finite. Without loss of generality we rank a real-valued function $f(\omega_k)$ on Ω as $|f(\omega_1)| \geq |f(\omega_2)| \geq ... \geq |f(\omega_n)|$ and $f(\omega_{n+1}) = 0$. Then, the expected value (i.e., Choquet-like aggregation) of f on Ω with respect to a bi-capacity η is expressed as

$$E_\eta[f] = \sum_{k=1}^{n} \big(|f(\omega_{(k)})| - |f(\omega_{(k+1)})|\big)\eta\big(\{\omega_{(1)}, ...\omega_{(k)}\} \cap N^+, \{\omega_{(1)}, ...\omega_{(k)}\} \cap N^-\big), \tag{A.2}$$

where $N^+ = \{\omega \in \Omega \mid f(\omega) \geq 0\}$ and $N^- = \{\omega \in \Omega \mid f(\omega) < 0\}$.

Example A.1 Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and the function f on Ω takes the values of $f(\omega_1) = 4$, $f(\omega_2) = 3$, and $f(\omega_3) = -2$. Then $N^+ = \{\omega_1, \omega_2\}$, $N^- = \{\omega_3\}$,

$$\begin{aligned} E_\eta[f] =& \big(|f(\omega_1)| - |f(\omega_2)|\big)\eta(\{\omega_1\}, \{\emptyset\}) \\ &+ \big(|f(\omega_2)| - |f(\omega_3)|\big)\eta(\{\omega_1, \omega_2\}, \{\emptyset\}) + |f(\omega_3)|\eta(\{\omega_1, \omega_2\}, \{\omega_3\}) \\ =& \eta(\{\omega_1\}, \{\emptyset\}) + \eta(\{\omega_1, \omega_2\}, \{\emptyset\}) + 2\eta(\{\omega_1, \omega_2\}, \{\omega_3\}). \end{aligned}$$

The mathematical treatment of bi-capacities and Choquet-like aggregation can be found in Grabisch and Labreuche (2005a, 2005b) and Labreuche and Grabisch (2006).

Fuzzy Set Theory. The ideas of fuzzy sets and fuzzy logic date back to Black (1937) and it has been mathematically formalized by Zadeh (1965). The most common type

---

[12] The same symbol, η, is used for both, capacities and bi-capacities. This should not create any notational confusion since $\eta(\cdot)$ is a capacity and $\eta(\cdot, \cdot)$ is a bi-capacity. A fuzzy version of a bi-capacity $\eta(\cdot, \cdot)$ is further denoted by a superimposed tilde, $\tilde{\eta}(\cdot, \cdot)$.

of fuzzy sets is called standard fuzzy sets. Each of the standard fuzzy sets is uniquely defined by a membership function of the form $\mu_{\tilde{A}} : \Omega \rightarrow [0, 1]$, where Ω denotes a universal set and $\tilde{A}$ is a fuzzy subset of Ω. Since a characteristic function of classical sets is a special case of a membership function of fuzzy sets, $\{0, 1\} \subseteq [0, 1]$, fuzzy sets are considered a formal generalization of classical sets.

The three basic operations on sets—complementation, intersection, and union—are not unique in fuzzy sets as they are in classical sets. The standard complement of a fuzzy set $\tilde{A}$ is a fuzzy set $\tilde{A}^c$ with the membership function $\mu_{\tilde{A}^c} = 1 - \mu_{\tilde{A}}$. The standard intersection of two fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a fuzzy set with the membership function $\mu_{\tilde{A}\cap\tilde{B}}(\omega) = \min\{\mu_{\tilde{A}}(\omega), \mu_{\tilde{B}}(\omega)\}$ and the standard union of two fuzzy sets is also a fuzzy set with $\mu_{\tilde{A}\cup\tilde{B}}(\omega) = \max\{\mu_{\tilde{A}}(\omega), \mu_{\tilde{B}}(\omega)\}$, where $\omega \in \Omega$ (see Bellman and Giertz (1973) for axiomatization of these standard operations). In addition, a fuzzy set $\tilde{A}$ is said to be a subset of fuzzy set $\tilde{B}$, $\tilde{A} \subseteq \tilde{B}$, if and only if $\mu_{\tilde{A}}(\omega) \leq \mu_{\tilde{B}}(\omega)$, $\forall\omega \in \Omega$ given that fuzzy sets $\tilde{A}$ and $\tilde{B}$ are defined on Ω.

One of the most important concepts of fuzzy sets is an α-cut of a fuzzy set which is one way of connecting fuzzy sets to classical sets. An α-cut of a fuzzy set $\tilde{A}$ on Ω denoted as ${}^{\alpha}A$ is a classical set that satisfies ${}^{\alpha}A = \{\omega \mid \mu_{\tilde{A}}(\omega) \geq \alpha\}$, where $\alpha \in [0, 1]$. A strong α-cut, denoted as ${}^{\alpha+}A$, is similar to the α-cut representation, ${}^{\alpha+}A = \{\omega \mid \mu_{\tilde{A}}(\omega) > \alpha\}$, but with a stronger condition. ${}^{0+}A$ and ${}^{1}A$ are called support and core of a fuzzy set $\tilde{A}$, respectively. When the core of a fuzzy set $\tilde{A}$ is not empty, ${}^{1}A \neq \emptyset$, $\tilde{A}$ is called normal, otherwise it is called subnormal. A fuzzy set is convex if and only if all its α-cuts are convex sets as in the classical sense.

Definition A.2 A fuzzy set $\tilde{A}$ on $\mathcal{R}$ (a set of real numbers) is a fuzzy number if (i) $\tilde{A}$ is a normal fuzzy set, (ii) ${}^{\alpha}A$ is a closed interval for every $\alpha \in (0, 1]$ and (iii) the support of $\tilde{A}$ is bounded.

When fuzzy numbers are used to formulate linguistic concepts such as very small, small, and so on, the final constructs are called linguistic variables.

Definition A.3 Let $\mathcal{E}^n$ be a space of all fuzzy subsets of $\mathcal{R}^n$ consisting of fuzzy sets which are normal, fuzzy convex, upper semi-continuous with compact support. A fuzzy function is a mapping from universal set Ω to $\mathcal{E}^n$, $\tilde{f} : \Omega \rightarrow \mathcal{E}^n$.

Definition A.4 Let $\tilde{A}, \tilde{B} \in \mathcal{E}^n$. If there exists $\tilde{C} \in \mathcal{E}^n$ such that $\tilde{A} = \tilde{B} + \tilde{C}$, then $\tilde{C}$ is called a Hukuhara difference $(-_h)$ of $\tilde{A}$ and $\tilde{B}$.

Example A.2 Let $\tilde{A}$ and $\tilde{B}$ be triangular fuzzy sets $\tilde{A} = (5, 7, 9)$ and $\tilde{B} = (1, 2, 3)$. Then, $\tilde{A} -_h \tilde{B} = (5, 7, 9) - (1, 2, 3) = (5 - 1, 7 - 2, 9 - 3) = (4, 5, 6)$. Hence, $\tilde{B} + (\tilde{A} -_h \tilde{B}) = (1, 2, 3) + (4, 5, 6) = (5, 7, 9) = \tilde{A}$.

Definition A.5 Given a fuzzy number $\tilde{A}$ on Ω, absolute value $|\tilde{A}|$ is defined as

$$\mu_{|\tilde{A}|}(\omega) = \begin{cases} \max\left(\mu_{\tilde{A}}(\omega), \mu_{-\tilde{A}}(\omega)\right) & \text{for } \omega \in \mathcal{R}^+, \\ 0 & \text{for } \omega \in \mathcal{R}^-. \end{cases} \tag{A.3}$$

Table 6.6 Fuzzy outcomes of each act under different states This table shows the precisiated gains or losses of bonds—$\tilde{f}_1$ (high income under $\tilde{s}_1$, medium income under $\tilde{s}_2$, less than medium income under $\tilde{s}_3$ and small income under $\tilde{s}_4$), stocks—$\tilde{f}_2$ (very high income under $\tilde{s}_1$, medium income under $\tilde{s}_2$, small income under $\tilde{s}_3$ and a notable loss under $\tilde{s}_4$), and term deposit—$\tilde{f}_3$ (the same medium income in all 4 fuzzy states of economy)

	States of the economy			
Acts	Strong growth (s_1)	Moderate growth (s_2)	Stable (s_3)	Recession (s_4)
Bonds ($\tilde{f}_1$)	(8, 11, 14)	(5, 8, 11)	(3, 6, 9)	(1, 3, 5)
Stocks ($\tilde{f}_2$)	(11, 15, 19)	(5, 8, 11)	(1, 3, 5)	(−3, −1.5, 0)
Term deposit ($\tilde{f}_3$)	(5, 8, 11)	(5, 8, 11)	(5, 8, 11)	(5, 8, 11)

Appendix B—Decision Theory 4

In this Appendix, we solve the decision problem outlined in Section 3.4 using behavioral decision-making with combined states under imperfect information (BDM) of Aliev et al. (2013). BDM combines fuzzy states of nature and fuzzy states of the decision maker as $\Omega = S \times H$ (cartesian product of S and H) with the elements of $\tilde{\omega}_i^j = (\tilde{s}_i, \tilde{h}_j)$ to account for the fundamental level dependence of S and H as shown by Kahneman and Tversky (1979). More precisely, Kahneman and Tversky (1979) show that a DM is risk-averse in the positive domain and risk-seeking in the negative domain. Neither classical measures nor capacities is adequate to capture the given imprecision induced by the natural language and the dependence of S and H. At this information class, among the fuzzy set of actions, $A = \{\tilde{f} \in A \mid \tilde{f} : \Omega \to X\}$ where X denotes a space of fuzzy outcomes, BDM determines an optimal action $\tilde{f}^* \in A$ as $\tilde{U}(\tilde{f}^*) = max_{\tilde{f} \in A} \int_{\Omega} \tilde{U}(\tilde{f}(\tilde{\omega})) \cdot d\tilde{\eta}(\cdot, \cdot)$, implying that an overall utility of an action is determined by a fuzzy number valued bi-capacity based aggregation over space Ω.

To solve the problem with BDM, suppose the outcomes at each fuzzy states of economy are represented by triangular fuzzy numbers given in Table 6.6. In other words, the fuzzy numbers in Table 6.6 are precisiated forms of the given linguistic outcomes.

Then we assign fuzzy utilities $\tilde{u}(\tilde{f}_k(\tilde{\omega}_i^j))$ (utility of action $\tilde{f}_k$ under state of economy $\tilde{s}_i$ when the DM's own state is h_j) by applying a technique of value function of Tversky and Kahneman (1992) as

$$\tilde{u}(\tilde{f}_k(\tilde{\omega}_i^1)) = \begin{cases} (\tilde{f}_k(\tilde{s}_i))^{\alpha} & \text{when } \tilde{f}_k(\tilde{s}_i) \geq 0, \\ -\lambda(-\tilde{f}_k(\tilde{s}_i))^{\beta} & \text{when } \tilde{f}_k(\tilde{s}_i) < 0; \end{cases} \tag{B.1}$$

$$\tilde{u}(\tilde{f}_k(\tilde{\omega}_i^2)) = \begin{cases} (\tilde{f}_k(\tilde{s}_i))^{\beta} & \text{when } \tilde{f}_k(\tilde{s}_i) \geq 0, \\ -\lambda(-\tilde{f}_k(\tilde{s}_i))^{\alpha} & \text{when } \tilde{f}_k(\tilde{s}_i) < 0; \end{cases} \tag{B.2}$$

Table 6.7 Fuzzy utilities under different states of economy and decision maker This table shows the absolute value of the utility of an action $\tilde{f}_k$ under state of economy $\tilde{s}_i$ when the DM's own state is h_j using Eqs. (B.1) and (B.2) with $\alpha = 0.88$, $\beta = 1.25$ and $\lambda = 2.25$. The utilities of each action under different combined states are reported in a descending order

$\tilde{u}(\tilde{f}_1(\tilde{\omega}_1^2)) \approx (13, 20, 27)$	$\tilde{u}(\tilde{f}_1(\tilde{\omega}_2^2)) \approx (7, 13, 20)$	$\tilde{u}(\tilde{f}_1(\tilde{\omega}_3^2)) \approx (4, 9, 16)$
$\tilde{u}(\tilde{f}_1(\tilde{\omega}_1^1)) \approx (6, 8, 10)$	$\tilde{u}(\tilde{f}_1(\tilde{\omega}_2^1)) \approx (4, 6, 8)$	$\tilde{u}(\tilde{f}_1(\tilde{\omega}_3^1)) \approx (3, 5, 7)$
$\tilde{u}(\tilde{f}_1(\tilde{\omega}_4^2)) \approx (1, 4, 7)$	$\tilde{u}(\tilde{f}_1(\tilde{\omega}_4^1)) \approx (1, 3, 4)$	
$\tilde{u}(\tilde{f}_2(\tilde{\omega}_1^2)) \approx (20, 30, 40)$	$\tilde{u}(\tilde{f}_2(\tilde{\omega}_2^2)) \approx (7, 13, 20)$	$\tilde{u}(\tilde{f}_2(\tilde{\omega}_1^1)) \approx (8, 11, 13)$
$\tilde{u}(\tilde{f}_2(\tilde{\omega}_2^1)) \approx (4, 6, 8)$	$\lvert\tilde{u}(\tilde{f}_2(\tilde{\omega}_4^1))\rvert \approx (0, 4, 9)$	$\tilde{u}(\tilde{f}_2(\tilde{\omega}_3^2)) \approx (1, 4, 7)$
$\lvert\tilde{u}(\tilde{f}_2(\tilde{\omega}_4^2))\rvert \approx (0, 3, 6)$	$\tilde{u}(\tilde{f}_2(\tilde{\omega}_3^1)) \approx (1, 3, 4)$	
$\tilde{u}(\tilde{f}_3(\tilde{\omega}_1^2)) \approx (7, 13, 20)$	$\tilde{u}(\tilde{f}_3(\tilde{\omega}_2^2)) \approx (7, 13, 20)$	$\tilde{u}(\tilde{f}_3(\tilde{\omega}_3^2)) \approx (7, 13, 20)$
$\tilde{u}(\tilde{f}_3(\tilde{\omega}_4^2)) \approx (7, 13, 20)$	$\tilde{u}(\tilde{f}_3(\tilde{\omega}_1^1)) \approx (4, 6, 8)$	$\tilde{u}(\tilde{f}_3(\tilde{\omega}_2^1)) \approx (4, 6, 8)$
$\tilde{u}(\tilde{f}_3(\tilde{\omega}_3^1)) \approx (4, 6, 8)$	$\tilde{u}(\tilde{f}_3(\tilde{\omega}_4^1)) \approx (4, 6, 8)$	

where $\alpha = 0.88$, $\beta = 1.25$ and $\lambda = 2.25$. For instance,

$$\tilde{u}(\tilde{f}_1(\tilde{\omega}_1^1)) = (\tilde{f}_1(\tilde{s}_1))^{\alpha} = (8^{0.88}, 11^{0.88}, 14^{0.88}) \approx (6, 8, 10), \tag{B.3}$$

$$\tilde{u}(\tilde{f}_1(\tilde{\omega}_1^2)) = (\tilde{f}_1(\tilde{s}_1))^{\beta} = (8^{1.25}, 11^{1.25}, 14^{1.25}) \approx (13, 20, 27). \tag{B.4}$$

Table 6.7 shows the absolute values of utilities in a descending order.

After assigning fuzzy utilities to each act, the next step is to construct a fuzzy joint probability distribution (FJP) $\tilde{P}$ on Ω given the fuzzy marginal probabilities of $S = \{\tilde{s}_1, \tilde{s}_2, \tilde{s}_3, \tilde{s}_4\}$ and $H = \{h_1, h_2\}$. With the given probabilities in natural language, we precisiate the fuzzy marginal probability distributions of S and H by the following triangular fuzzy numbers.[13]

$$\begin{gathered}\tilde{P}(\tilde{s}_1) = (0.45, 0.50, 0.55), \quad \tilde{P}(\tilde{s}_2) = (0.325, 0.35, 0.375), \\ \tilde{P}(\tilde{s}_3) = (0.1, 0.125, 0.15), \quad \tilde{P}(\tilde{s}_4) = (0, 0.025, 0.125) \text{ (computed)}, \\ \tilde{P}(h_1) = (0.65, 0.70, 0.75), \quad \tilde{P}(h_2) = (0.25, 0.30, 0.35) \text{ (computed)}.\end{gathered} \tag{B.5}$$

Given the fuzzy marginal probability distributions of S and H, we obtain the FJP distribution on the base of positive and negative dependence concept of Wise and Henrion (1985).[14] Formally, the FJP is obtained following

[13] By convention, we precisiate $(n-1)$ of the given linguistic probabilities and compute the last one in order to add up total probabilities to 1.

[14] Given the numerical probabilities $P(A)$ and $P(B)$, the joint probability of A and B is $P(A, B) = P(A) \cdot P(B)$ if A and B are independent, $P(A, B) = \min\big(P(A), P(B)\big)$ if A and B are positively dependent, and $P(A, B) = \max\big(P(A) + P(B) - 1, 0\big)$ if A and B have negative dependence. Eqs. (B.6) and (B.7) are the extensions of these formulations to fuzzy probabilities via α-cuts.

$$\tilde{p}(\tilde{s}_i, h_j) = \bigcup_{\alpha\in[0,1]} \alpha\big[{}^{\alpha}p_1(s_i){}^{\alpha}p_1(h_j), \min\big({}^{\alpha}p_2(s_i), {}^{\alpha}p_2(h_j)\big)\big], \tag{B.6}$$

$$\tilde{p}(\tilde{s}_i, h_j) = \bigcup_{\alpha\in[0,1]} \alpha\big[\max\big({}^{\alpha}p_1(s_i) + {}^{\alpha}p_1(h_j) - 1, 0\big), {}^{\alpha}p_2(s_i){}^{\alpha}p_2(h_j)\big] \tag{B.7}$$

for positive and negative dependence, respectively. For $\tilde{f}_1$ and $\tilde{f}_3$, there are positive dependences between, $(\tilde{s}_1, h_1)$, $(\tilde{s}_2, h_1)$, $(\tilde{s}_3, h_1)$, $(\tilde{s}_4, h_1)$ and negative dependences between $(\tilde{s}_1, h_2)$, $(\tilde{s}_2, h_2)$, $(\tilde{s}_3, h_2)$, $(\tilde{s}_4, h_2)$. For $\tilde{f}_2$ there are positive dependences between, $(\tilde{s}_1, h_1)$, $(\tilde{s}_2, h_1)$, $(\tilde{s}_3, h_1)$, $(\tilde{s}_4, h_2)$ and negative dependences between $(\tilde{s}_1, h_2)$, $(\tilde{s}_2, h_2)$, $(\tilde{s}_3, h_2)$, $(\tilde{s}_4, h_1)$. This is because, according to Kahneman and Tversky (1979), people are risk-averse in the positive domain and risk-seeking in the negative domain. Then, we compute $\tilde{p}(\tilde{s}_1, h_1)$ for $\tilde{f}_1$, $\tilde{f}_2$ and $\tilde{f}_3$ given $\alpha = 0, 0.5, 1$ as

$$\begin{aligned}\big[{}^{0}p_1(\tilde{s}_1){}^{0}p_1(h_1), \min\big({}^{0}p_2(\tilde{s}_1){}^{0}p_2(h_1)\big)\big] &= [0.45 \cdot 0.65, \min(0.55, 0.75)] \\ &\approx [0.293, 0.55];\end{aligned} \tag{B.8}$$

$$\begin{aligned}\big[{}^{.5}p_1(\tilde{s}_1){}^{.5}p_1(h_1), \min\big({}^{.5}p_2(\tilde{s}_1){}^{.5}p_2(h_1)\big)\big] &= [0.475 \cdot 0.675, \min(0.525, 0.725)] \\ &= [0.32, 0.525];\end{aligned} \tag{B.9}$$

$$\begin{aligned}\big[{}^{1}p_1(\tilde{s}_1){}^{1}p_1(h_1), \min\big({}^{1}p_2(\tilde{s}_1){}^{1}p_2(h_1)\big)\big] &= [0.5 \cdot 0.7, \min(0.5, 0.7)] \\ &= [0.35, 0.5].\end{aligned} \tag{B.10}$$

Hence, $\tilde{p}(\tilde{s}_1, h_1)$ can be approximated by $(0.293, 0.35, 0.5, 0.55)$ trapezoidal fuzzy number. Following these steps, the FJPs of $\tilde{s}_i$ and h_j for $\tilde{f}_1$ and $\tilde{f}_3$, respectively, are

$$\begin{aligned}&\tilde{p}(\tilde{s}_1, h_1) = (0.293, 0.35, 0.5, 0.55),\ \tilde{p}(\tilde{s}_2, h_1) = (0.211, 0.245, 0.350, 0.375), \\ &\tilde{p}(\tilde{s}_3, h_1) = (0.065, 0.088, 0.125, 0.15),\ \tilde{p}(\tilde{s}_4, h_1) = (0, 0.018, 0.025, 0.125),\end{aligned} \tag{B.11}$$

$$\begin{aligned}&\tilde{p}(\tilde{s}_1, h_2) = (0, 0, 0.150, 0.193),\ \tilde{p}(\tilde{s}_2, h_2) = (0, 0, 0.105, 0.131), \\ &\tilde{p}(\tilde{s}_3, h_2) = (0, 0, 0.038, 0.053),\ \tilde{p}(\tilde{s}_4, h_2) = (0, 0, 0.008, 0.044).\end{aligned} \tag{B.12}$$

The FJPs for $\tilde{f}_2$ are the same as $\tilde{f}_1$ and $\tilde{f}_3$ for all combinations except

$$\tilde{p}(\tilde{s}_4, h_1) = (0, 0, 0.008, 0.044),\ \tilde{p}(\tilde{s}_4, h_2) = (0, 0.018, 0.025, 0.125), \tag{B.13}$$

due to an inverse relationship in $\tilde{s}_4$.

The next step is to construct a fuzzy valued bi-capacity $\tilde{\eta}(\cdot,\cdot)$ based on the obtained FJPs. A fuzzy valued bi-capacity is defined, $\tilde{\eta}(\tilde{A}, \tilde{B}) = \tilde{\eta}(\tilde{A}) - \tilde{\eta}(\tilde{B})$, as a difference of fuzzy-valued lower probabilities $\tilde{\eta}(\tilde{A})$ and $\tilde{\eta}(\tilde{B})$. Given a set

$\Omega = \{\omega_1^1, \omega_1^2, \omega_2^1, ..., \omega_4^1, \omega_4^2\}$ and its power set $\mathcal{F}(\Omega)$, let ${}^{\alpha}I = \{[{}^{\alpha}l_i, {}^{\alpha}u_i] \mid i \in \mathbb{N}_8\}$ denote 8-tuples of probability intervals on $\omega_i^j \in \Omega$ where ${}^{\alpha}l_i$ and ${}^{\alpha}u_i$ denote corresponding lower and upper bounds of α-cuts of the computed FJPs, respectively. Consistent with decision theories 2 and 3, let $\tilde{\mathcal{M}}$ denote a set of fuzzy probabilities $\tilde{p}$ on $\mathcal{F}(\Omega)$ satisfying

$$\tilde{\mathcal{M}} = \{\tilde{p} \mid {}^{\alpha}l(\omega_i^j) \leq {}^{\alpha}p(\omega_i^j) \leq {}^{\alpha}u(\omega_i^j), i \in \mathbb{N}_4, j \in \mathbb{N}_2, \sum_{\omega_i^j \in \Omega} \tilde{p}(\omega_i^j) = 1\}. \quad \text{(B.14)}$$

From the fuzzy probabilities in set $\tilde{\mathcal{M}}$, the lower probability measure is defined for all $\tilde{A} \in \mathcal{F}(\Omega)$ as $\tilde{\eta}(\tilde{A}) = \inf_{\tilde{p} \in \tilde{\mathcal{M}}} \sum_{x_i \in \tilde{A}} \tilde{p}(x_i)$. An α-cut of a fuzzy lower probability measure, ${}^{\alpha}\eta$, are calculated as in decision theory 2.

$$ {}^{\alpha}\eta(\tilde{A}) = \max \left\{ \sum_{x_i \in \tilde{A}} {}^{\alpha}l(x_i), 1 - \sum_{x_i \notin \tilde{A}} {}^{\alpha}u(x_i) \right\}, \forall \tilde{A} \in \mathcal{F}(\Omega). \quad \text{(B.15)}$$

For clarity, let us exemplify;

$$ {}^{\alpha}\eta(\tilde{\omega}_1^2) = \max\{{}^{\alpha}l(\tilde{\omega}_1^2), 1 - \sum_{i \neq 1, j \neq 2} {}^{\alpha}u(\tilde{\omega}_i^j)\} = {}^{\alpha}l(\tilde{\omega}_1^2) = 0. \quad \text{(B.16)}$$

Hence, $\eta(\tilde{\omega}_1^2) = (0, 0, 0)$.

$$\begin{aligned} {}^{\alpha}\eta(\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1) &= \max \Big\{ \big({}^{\alpha}l(\tilde{\omega}_1^2) + {}^{\alpha}l(\tilde{\omega}_2^2) + {}^{\alpha}l(\tilde{\omega}_1^1) + {}^{\alpha}l(\tilde{\omega}_2^1)\big), \\ &\qquad \big(1 - {}^{\alpha}u(\tilde{\omega}_3^1) - {}^{\alpha}u(\tilde{\omega}_3^2) - {}^{\alpha}u(\tilde{\omega}_4^1) - {}^{\alpha}u(\tilde{\omega}_4^2)\big) \Big\} \\ &= \big(1 - {}^{\alpha}u(\tilde{\omega}_3^1) - {}^{\alpha}u(\tilde{\omega}_3^2) - {}^{\alpha}u(\tilde{\omega}_4^1) - {}^{\alpha}u(\tilde{\omega}_4^2)\big) = 0.629 + 0.176 \cdot \alpha. \end{aligned} \quad \text{(B.17)}$$

Hence, $\eta(\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1) \approx (0.63, 0.8, 0.8)$. Table 6.8 reports the values of capacities and bi-capacities for all states of f_2.

Finally, we calculate fuzzy overall utilities by a fuzzy-valued bi-capacity based aggregation over space Ω using generalized version of Choquet-like aggregation. For $\tilde{f}_2$ the overall utilities are given by

$$\begin{aligned} \tilde{U}(\tilde{f}_2) = &\big(|\tilde{u}(\tilde{f}_2(\tilde{\omega}_1^2))| -_h |\tilde{u}(\tilde{f}_2(\tilde{\omega}_2^2))|\big)\tilde{\eta}\big(\{\tilde{\omega}_1^2\}, \{\emptyset\}\big) \\ &+ \big(|\tilde{u}(\tilde{f}_2(\tilde{\omega}_2^2))| -_h |\tilde{u}(\tilde{f}_2(\tilde{\omega}_1^1))|\big)\tilde{\eta}\big(\{\tilde{\omega}_1^2, \tilde{\omega}_2^2\}, \{\emptyset\}\big) \\ &+ \big(|\tilde{u}(\tilde{f}_2(\tilde{\omega}_1^1))| -_h |\tilde{u}(\tilde{f}_2(\tilde{\omega}_2^1))|\big)\tilde{\eta}\big(\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1\}, \{\emptyset\}\big) \\ &+ \big(|\tilde{u}(\tilde{f}_2(\tilde{\omega}_2^1))| -_h |\tilde{u}(\tilde{f}_2(\tilde{\omega}_4^1))|\big)\tilde{\eta}\big(\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1\}, \{\emptyset\}\big) \\ &+ \big(|\tilde{u}(\tilde{f}_2(\tilde{\omega}_4^1))| -_h |\tilde{u}(\tilde{f}_2(\tilde{\omega}_3^2))|\big)\tilde{\eta}\big(\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1\}, \{\tilde{\omega}_4^1\}\big) \end{aligned}$$

Table 6.8 Fuzzy-valued bi-capacities for $\tilde{f}_2$ This table reports the fuzzy-valued bi-capacities for each combined state of f_2. The first column shows the combined states, the second column shows the first element of the bi-capacity ($\tilde{\eta}(A)$), and the third column shows the second element of the bi-capacity ($\tilde{\eta}(B)$). The last column reports the bi-capacity of a combined state as the difference of the first and second element ($\tilde{\eta}(A) - \tilde{\eta}(B)$)

$A, B \subset \Omega$	$\tilde{\eta}(A)$	$\tilde{\eta}(B)$	$\tilde{\eta}(A, B)$
$\{\tilde{\omega}_1^2\}, \{\emptyset\}$	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
$\{\tilde{\omega}_1^2, \tilde{\omega}_2^2\}, \{\emptyset\}$	(0,0,0)	(0,0,0)	(0,0,0)
$\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1\}, \{\emptyset\}$	(0.29, 0.45, 0.45)	(0, 0, 0)	(0.29, 0.45, 0.45)
$\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1\}, \{\emptyset\}$	(0.63, 0.8, 0.8)	(0, 0, 0)	(0.63, 0.8, 0.8)
$\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1\}, \{\tilde{\omega}_4^1\}$	(0.63, 0.8, 0.8)	(0, 0, 0)	(0.63, 0.8, 0.8)
$\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1, \tilde{\omega}_3^2\}, \{\tilde{\omega}_4^1\}$	(0.68, 0.84, 0.84)	(0, 0, 0)	(0.68, 0.84, 0.84)
$\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1, \tilde{\omega}_3^2\}, \{\tilde{\omega}_4^1, \tilde{\omega}_4^2\}$	(0.68, 0.84, 0.84)	(0, 0.02, 0.02)	(0.68, 0.82, 0.82)
$\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1, \tilde{\omega}_3^2, \tilde{\omega}_3^1\}, \{\tilde{\omega}_4^1, \tilde{\omega}_4^2\}$	(0.83,0.97,0.97)	(0, 0.02, 0.02)	(0.83, 0.95, 0.95)

$$
\begin{aligned}
&+ \big(|\tilde{u}(\tilde{f}_2(\tilde{\omega}_3^2))| -_h |\tilde{u}(\tilde{f}_2(\tilde{\omega}_4^2))|\big)\tilde{\eta}\big(\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1, \tilde{\omega}_3^2\}, \{\tilde{\omega}_4^1\}\big) \\
&+ \big(|\tilde{u}(\tilde{f}_2(\tilde{\omega}_4^2))| -_h |\tilde{u}(\tilde{f}_2(\tilde{\omega}_3^1))|\big)\tilde{\eta}\big(\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1, \tilde{\omega}_3^2, \}, \{\tilde{\omega}_4^1, \tilde{\omega}_4^2\}\big) \\
&+ \big(|\tilde{u}(\tilde{f}_2(\tilde{\omega}_3^1))|\big)\tilde{\eta}\big(\{\tilde{\omega}_1^2, \tilde{\omega}_2^2, \tilde{\omega}_1^1, \tilde{\omega}_2^1, \tilde{\omega}_3^2\}, \{\tilde{\omega}_4^1, \tilde{\omega}_4^2\}\big) \\
=&(3.99, 7.49, 9.61)
\end{aligned}
\tag{B.18}
$$

where $-_h$ is a Hukuhara difference (see Appendix A). The values of utilities for other alternatives similarly follow as $\tilde{U}(\tilde{f}_1) = (4, 6.92, 8.91)$ and $\tilde{U}(\tilde{f}_3) = (4.12, 6.23, 8.25)$. Finally, we rank these fuzzy numbers as $\tilde{f}_2 \succ \tilde{f}_1 \succ \tilde{f}_3$.

References

Aliev R, Alizadeh A, Huseynov O (2015) The arithmetic of discrete z-numbers. Inf Sci 290:134–155

Aliev R, Pedrycz W, Fazlollahi B, Huseynov O, Alizadeh AV, Guirimov B (2012) Fuzzy logic-based generalized decision theory with imperfect information. Inf Sci 189:18–42

Aliev R, Pedrycz W, Huseynov O (2013) Behavioral decision making with combined states under imperfect information. Int J Inf Technol Decis Making 12(03):619–645

Aliev R, Pedrycz W, Kreinovich V, Huseynov O (2016) The general theory of decisions. Inf Sci 327:125–148

Aliyev N, He X (2022a) Ambiguity and information tradeoffs. Working paper, UTS Business School

Aliyev N, He X (2022b) Ambiguous price formation. Working paper, UTS Business School

Aliyev N, He X, Putnins T (2022) Learning about adverse selection in markets. Working paper, UTS Business School

Arrow KJ (1951) Alternative approaches to the theory of choice in risk-taking situations. Econometrica 19(4):404–437

Bachelier L (1900) Théorie de la spéculation. Gauthier-Villars

Banerjee S, Green B (2015) Signal or noise? uncertainty and learning about whether other traders are informed. J Financ Econ 117(2):398–423
Bellman R, Giertz M (1973) On the analytic formalism of the theory of fuzzy sets. Inf Sci 5:149–156
Bewley TF (2002) Knightian decision theory. Part I. Decisions Econ Finan 25(2):79–110
Black M (1937) Vagueness. an exercise in logical analysis. Philos Sci 4(4):427–455
Bondt WFM, Thaler R (1985) Does the stock market overreact? J Financ 40(3):793–805
Choquet G (1955) Theory of capacities. Ann Inst Fourier (Grenoble) 5:131–295
Dempster AP (1967) Upper and lower probabilities induced by a multivalued mapping. Ann Math Stat 38(2):325–339
Dow J, Werlang SR (1992) Uncertainty aversion, risk aversion, and the optimal choice of portfolio. Econometrica 60(1):197–204
Easley D, O'Hara M, Yang L (2013) Opaque trading, disclosure, and asset prices: implications for hedge fund regulation. Rev Financ Stud 27(4):1190–1237
Ellsberg D (1961) Risk, ambiguity, and the savage axioms. Quart J Econ 643–669
Epstein LG, Schneider M (2008) Ambiguity, information quality, and asset pricing. J Financ 63(1):197–228
Fama E (1965) The behavior of stock-market prices. J Bus 38(1):34–105
Friedman M (1953) The case for flexible exchange rates. University of Chicago Press
Friedman M, Savage LJ (1948) The utility analysis of choices involving risk. J Polit Econ 56(4):279–304
Ghirardato P, Maccheroni F, Marinacci M (2004) Differentiating ambiguity and ambiguity attitude. J Econ Theory 118(2):133–173
Gilboa I, Maccheroni F, Marinacci M, Schmeidler D (2010) Objective and subjective rationality in a multiple prior model. Econometrica 78(2):755–770
Gilboa I, Schmeidler D (1989) Maxmin expected utility with non-unique prior. J Math Econ 18(2):141–153
Grabisch M, Labreuche C (2005) Bi-capacities i: Definition, Möbius transform and interaction. Fuzzy Sets Syst 151(2):211–236
Grabisch M, Labreuche C (2005) Bi-capacities ii: the Choquet integral. Fuzzy Sets Syst 151(2):237–259
Grossman SJ, Stiglitz JE (1980) On the impossibility of informationally efficient markets. Am Econ Rev 70(3):393–408
Gul F, Pesendorfer W (2014) Expected uncertain utility theory. Econometrica 82(1):1–39
Ju N, Miao J (2012) Ambiguity, learning, and asset returns. Econometrica 80(2):559–591
Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47(2):263–291
Keynes JM (1921) A treatise on probability. Cambridge University Press
Klibanoff P, Marinacci M, Mukerji S (2005) A smooth model of decision making under ambiguity. Econometrica 73(6):1849–1892
Klir GJ (2005) Uncertainty and information: foundations of generalized information theory. Wiley
Knight FH (1921) Risk uncertainty and profit. Houghton Mifflin, Boston
Labreuche C, Grabisch M (2006) Generalized Choquet-like aggregation functions for handling bipolar scales. Eur J Oper Res 172(3):931–955
Lintner J (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Rev Econ Stat 47(1):13–37
Lo AW (2004) The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. J Portf Manag 30(5):15–29
Mandelbrot B (1963) The variation of certain speculative prices. J Bus 36(4):394–419
Markowitz H (1952) Portfolio selection. J Financ 7(1):77–91
Markowitz H (1952) The utility of wealth. J Polit Econ 60(2):151–158
Mehra R, Prescott EC (1985) The equity premium: a puzzle. J Monet Econ 15(2):145–161
Mossin J (1966) Equilibrium in a capital asset market. Econometrica 34(4):768–783
Osborne MFM (1959) Brownian motion in the stock market. Oper Res 7(2):145–173

Peters EE (1996) Chaos and order in the capital markets: a new view of cycles, prices, and market volatility. Wiley

Romer D (1993) Rational asset-price movements without news. Am Econ Rev 83(5):1112–30

Rubinstein M (2001) Rational markets: yes or no? the affirmative case. Financ Anal J 57(3):15–29

Samuelson PA (1965) Proof that properly anticipated prices fluctuate randomly. Ind Manag Rev 6(2):41–49

Savage LJ (1954) The foundations of statistics. Wiley, New York [2nd ed. 1972 (Dover, New York)]

Schmeidler D (1986) Integral representation without additivity. Proc Am Math Soc 97(2):255–261

Schmeidler D (1989) Subjective probability and expected utility without additivity. Econometrica 57(3):571–587

Shackle G (1949) Expectation in economics. Gibson Press, London [2nd ed. 1952 (Cambridge University Press)]

Shafer G (1976) A mathematical theory of evidence, vol 1. Princeton University Press, Princeton

Shafer G (1987) Probability judgment in artificial intelligence and expert systems. Stat Sci 2(1):3–16

Sharpe WF (1964) Capital asset prices: a theory of market equilibrium under conditions of risk. J Financ 19(3):425–442

Shefrin H (2000) Beyond greed and fear: understanding behavioral finance and the psychology of investing. Oxford University Press

Shiller RJ (1979) The volatility of long-term interest rates and expectations models of the term structure. J Polit Econ 87(6):1190–1219

Shiller RJ (1981) Do stock prices move too much to be justified by subsequent changes in dividends? Am Econ Rev 71(3):421–436

Thaler R (1980) Toward a positive theory of consumer choice. J Econ Behav Organ 1(1):39–60

Tversky A, Kahneman D (1992) Advances in prospect theory: cumulative representation of uncertainty. J Risk Uncertain 5(4):297–323

Von Neumann J, Morgenstern O (1944) Theory of games and economic behavior. Princeton University Press

Wise BP, Henrion M (1985) A framework for comparing uncertain inference systems to probability. In: Proceedings of the annual conference on uncertainty in artificial intelligence (UAI-85), vol 1, pp 99–108

Zadeh LA (1965) Fuzzy sets. Inf. Control 8(3):338–353

Zadeh LA (1978) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Syst 1(1):3–28

Zadeh LA (2011) A note on Z-numbers. Inf Sci 181(14):2923–2932

Chapter 7
Sand Castles and Financial Systems: A Sandpile Metaphor

Francesco Luna

Abstract Pooling financial resources will most likely reduce the number of financial distress episodes for individual institutions, but it may render the fewer crises significantly more severe for the system as a whole. This simulation exercise is designed to support this commonsensical intuition. The question we ask here is whether there is a level of interconnection among financial institutions that will maximize the welfare of the system by minimizing the average loss per institution per period of time. The metaphor is that of the sand pile. Following a lead suggested by Scott Page many years ago, the question is reformulated to assess the effect of placing the falling sand grains (financial shocks) to the location in the pile that is less likely to collapse: the result is a "sand castle". Is this construction advantageous for the system if the castle eventually collapses? Should the castle have some optimal structure? Does the optimal structure depend on the policy regime?.

Keywords Networks · Financial systems · Welfare losses · NPLs · Bankruptcies

7.1 Prelude

I met Professor Chen at UCLA when we both were studying to obtain our PhD in the department of economics. Like me, he was a member of the Vela Velupillai's club. Only independent and original thinkers were welcome and Shu (as we all referred to Professor Chen at the time) was one of the most adventurous intellects in the group. I believe he was the one to introduce me to the concept of genetic algorithm applied to economics. By the time I had digested the idea and applied it myself to some highly non-linear landscape, he moved to genetic programming and left me behind forever.

The simulation code was written in python and the notebooks are available upon requests from the author.

F. Luna (✉)
Office of Innovation and Change, International Monetary Fund, 700 19th Street, Washington DC, NW 20431, USA
e-mail: fluna@imf.org

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_7

Naturally, as all true adventurers, he was also very solidly rooted in tradition: he was the scout who led the group to one of the most authentic Chinese cuisine restaurants in Los Angeles. That is where the club's members would convene for the most interesting and inspiring conversations after Professor Velupillai's evening lectures. Discussions on coupled oscillators, Lorenz attractors, Turing machines, incompleteness, undecidability versus uncomputability, were all chewed up along superb mù xū ròu of the like I have never had again…and I am not only talking about the intellectual input.

Shu's smile and discreet sense of humor was proverbial. I will never forget when I eventually filed my dissertation and completed my Ph.D. He was among the first ones to reach out to me with his unmistakable warmth and wit. His message began: CongraDulations! So it is only fair now (and much more deserved) for me to join this group of friends, colleagues, former students, and fans to add my heart-felt "Congratulations" for such a brilliant, always engaged at the frontier, and extremely BUSY and TIRING career. Hence, Shu, let me wish you "all the Rest" for your future endeavors!

7.2 Introduction

One way to read the history of the legislation concerning the banking and financial system in the US could be the continuous attempt to reduce crises, but with the surprising effect of loosening the regulation originally designed to keep potential bankruptcies at the local level. The path seems to have been from free banking to central banking with the progressive imposition of the green back as the only legal currency, but also from very strict limitations on branching to open interstate banking. It is sufficient to recall that still in 1980 only the state of Maine allowed banks from other states to acquire Maine banks. It was only in 1994 with the Riegle-Neal Act that the McFadden Act (in effect since 1927!) was repealed. Of course, major financial crises occurred also during the McFadden act period, like the Savings and Loan crisis of the second half of the 80s and first half of the 90s.

The advantages deriving from the pooling of resources and diversification of assets are the universal maxims for solid banking practices, but, still, the idea of keeping the financial business to a local level has its charm: a crisis would be necessarily smaller with a much reduced risk of systemic effects. Were the legislators of a hundred years ago so wrong?

This paper is based on a metaphor: the sand pile. The simulations are built on a simplification of the original Abelian[1] sandpile model. Although the idea was proposed in various contexts (see Holroyd et al 2008), the model is often referred to as the Bak–Tang–Wiesenfeld model (Bak et al. 1987) and proposes a dynamic

[1] The name, as suggested by a reviewer, could be linked to the Norwegian mathematician Niels Henrik Abel probably for his work on group theory. So far, however, I must confess I have not been able to confirm this hypothesis.

framework exhibiting self-organized criticality. It could be referred to as an example of catastrophe theory.

However, I introduce a twist.[2] A sand pile could be modeled to grow into a sandcastle if the grains of sand are added with care. Will the castle be able to endure more shocks than a simple sand dune? And when it collapses, will the slide be broader? Maybe wider than the sum of all the slides that would have occurred if the pile had not been tampered with?

The results I will present are mostly qualitative, but calibration of various parameters and specific structures with weights assigned to links could be implemented to represent real financial sector networks.

In short, what I find is that the financial system structure will adjust differently to optimize social welfare depending on the intermediate goals of the authority: try to avoid a crisis or strictly implement the *laissez faire* principle. The optimization process in such a non-linear landscape is performed in two different ways: with a genetic algorithm and with a purely random search. Although quite rough, the GA mechanism consistently leads to superior results translating into a significant reduction in public welfare loss.[3]

7.3 The Model

The model used for the simulations in this paper is a simplification of the original Abelian sandpile model. Although the idea was proposed in various contexts (see Holroyd et al 2008), the model is often referred to as the Bak–Tang–Wiesenfeld model (Bak, et al. 1987) and proposes a dynamic framework exhibiting self-organized criticality. For the purpose of this exercise, rather than operating in a two-dimensional grid, I set it up on a network or, more precisely, on a directed graph.

I interpret the graph as a very simple financial system: each node is a deposit-taking financial institution, a bank. At each point in time, a random number of shocks hit the system. These shocks are new non-performing loans (NPLs) and they are the equivalent of the grains of sand falling onto the pile. Each bank can count on a set of branches or associates and partners. Each bank can endure a certain amount of non-performing loans (NPLs). Following the sand pile metaphor, once the amount

---

[2] I give the source of this "inspiration" in the next section as I cannot take all the blame or merit as the case may be. However, I now have learned thanks to a friendly reviewer, that such twist has been studied and classified in the combinatorics literature as "chip-firing games" on a graph. See for example Spencer (1986) and Bjorner et al. (1991)

[3] Since the purpose of this exercise was that of finding a possible link between the shape of the network and its performance, there is no formal investigation of the efficiency or effectiveness of the search algorithms employed. I have also considered another search mechanism, which I could call "gradient" as it looks for improvement by changing only one node. If the performance of this marginally modified graph is superior, the net becomes the basis for the next change and so on. I could not perfect the code for this algorithm, so the results are not reported here.

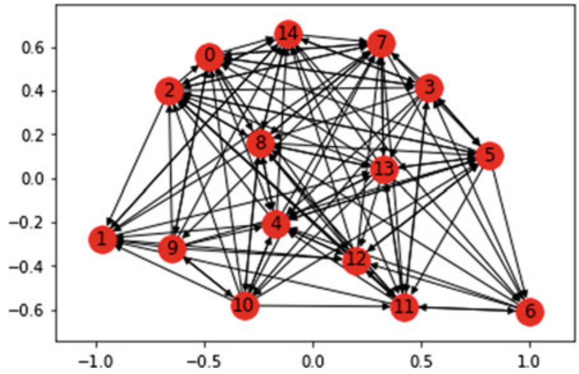


Fig. 7.1 Generic financial system as directed graph

of NPLs has reached a particular threshold, a slide occurs: part of the NPLs are transferred to the bank's associates (its "descendants" in the directed graph).

For this exploration, I have fixed the dimension of the directed graph at 15 nodes and a typical network would look like the one depicted in Fig. 7.1.

The graph is not fully connected: the number of directed links in and out of each node is randomly determined when the network is created. This implies that power relations are not symmetric and not necessarily each bank can "impose" to share a loss with its "ascendants". In Fig. 7.1, the symmetric relations (when they exist) are represented by lines with arrow points at both extremes like the one linking node 6 ad 11 (on the lower right corner). On the other hand, the link between 10 and 11 is only unidirectional: from 10 to 11.

The directed network is built by generating an asymmetric square matrix (15 × 15) with zeros along the diagonal (to avoid the paradox of a bank transferring a liability to itself). The following matrix lies behind the directed graph depicted in Fig. 7.1. It won't be a surprise that such a representation is very convenient. It is the perfect set up to apply a simple genetic algorithm, that will split and recombine the sequence of zeros and ones…

Hence, in Fig. 7.2, row n collects the "address" of all descendants of the nth node. The first row (indexed as zero in python), links node 0 to nodes 3, 4, 7, 8, 9, 12, and 14.

By construction, the threshold that will trigger an event (a slide) is set at 4 (this number seems very common among the simulations found on the web; see for example http://www.natureincode.com/code/various/sandpile.html). Each shock is unitary so a node (bank) that is hit by more than 4 shocks, will discharge exactly 4 NPLs to its descendants in the network: its NPL stock will be reduced by 4 and the bank will be tagged as "under distress".[4]

If the node has exactly four connections (descendants), each of these will receive one NPL to be added to their existing stock. If there are more than 4 connections,

[4] The status "under distress" will last for one day. Since the objective is to gauge the characteristics of a specific network, the same network will be used on each day in the life of a generation. The structure itself can be modified only across generations, as I will discuss later.

0	0	0	1	1	0	0	1	1	1	0	0	1	0	1
0	0	0	0	0	0	0	1	0	0	0	0	1	1	1
1	1	0	1	0	1	0	1	1	0	0	1	0	0	1
1	0	0	0	1	1	0	1	1	0	1	1	0	1	1
0	1	1	0	0	1	0	1	1	0	1	0	1	1	1
1	0	1	1	1	0	1	0	0	0	0	1	1	1	0
0	0	0	1	1	0	0	0	1	0	0	1	1	0	0
1	0	1	0	0	1	1	0	1	1	0	1	0	1	0
1	1	1	0	1	1	0	1	0	0	1	1	0	1	1
1	1	1	0	1	1	0	0	0	0	1	0	1	0	1
1	1	1	0	1	1	0	1	1	1	0	1	0	0	0
0	1	1	1	1	0	1	0	1	0	0	0	1	1	0
0	0	1	0	1	1	0	1	1	0	0	1	0	0	1
1	0	1	0	1	1	1	1	1	0	1	1	1	0	1
0	1	1	1	0	0	1	1	1	0	0	1	0	0	0

Fig. 7.2 Matrix representation of the financial system network

only four descendants will be chosen randomly to absorb the transfer. In both cases the "slide" is absorbed by the system and there is no explicit welfare loss.

In other words, I interpret as welfare loss only those losses that cannot be absorbed by the network and thus explicitly fall on the public. Such an event will occur when a node does not have at least 4 descendants. This may happen by construction of the original network (remember that the links are random so it may well be that a node does not have 4 descendants) and also because a bank that has already passed its threshold and is tagged as under distress, is isolated from its ascendants: all "incoming links" are broken for that period.

The simulations are organized on a set of 300 iterations (they could be interpreted as days) and the performance of the network is calculated at the end of the "year." The performance focuses on three dimensions: (i) the average time (number of days) between crisis events, (ii) the average number of slides that characterize the crisis events, and (iii) the average welfare loss generated by each crisis event. A crisis event is one that causes a slide, but this may not translate into a public welfare loss if the slide is absorbed by the specific node's descendants.

7.4 Public Intervention

In the late 90s as member of the Center for Computable Economics at UCLA, I was invited to participate in a research workshop at the Santa Fe' Institute. In that occasion, I was exposed for the first time to the sand pile model by Scott Page. In his presentation, professor Page talked about a few experiments that his group had run on the sand pile model with the following twist: the sand grains, rather than falling randomly on the surface of the pile, were artificially addressed to the spot where the probability of a slide was the smallest at that time. According to his account, the simulations showed that the slides were rarer, but more pronounced when they eventually occurred.

That was the inspiration for the current exercise. The macroeconomic literature abounds of studies that assess the impact of financial crises and measure the cost of such crises well beyond the immediate cost born by the shareholders. Credit is often seen almost like a public good; hence, the perception is that financial institutions' bankruptcies should be avoided whenever possible. This often requires the intervention of some authority that, through moral suasion and sometimes explicit financial support, facilitate mergers, acquisitions and recapitalizations.

In our set up, quite crudely, these actions can be represented by allocating financial shocks to the currently more robust nodes in the network. In this vein, the first set of simulations is designed both to check whether these interventions result in a reduction of the frequency of socially relevant crises and to identify possible side effects; in particular whether these rarer events are more costly. Over a longer period of time, the following exercise records the average time to a crisis, the average number of slides when the crisis hits (like in a domino effect), and the average cost incurred in a crisis. These series are collected under the two scenarios: with and without external intervention. The same set of shocks is imposed on a group of 50 randomly generated networks (Fig. 7.3).

According to these simulations, interventions as described above, reduced the frequency of crises, but the average cost of each of these rarer crises was higher. The conclusion is that interventions trying to reduce "localized" crises lead, on average, to a deeper social welfare per outcome.

7.5 Structure and Performance of the Financial System

This section focuses on the policy implications. The first lesson from the initial set of simulations seems to be: "do not tamper" with the unfettered functioning of the market. This "maxim" was certainly my prior and motivation to set up this model. However, by looking at the data with attention, I noticed a surprising result (for me, of course): the **total** social welfare **cost** imposed over the full period is **smaller in the "managed" case** ("tot Loss man" in the graph) (Fig. 7.4).

Apart from this surprising lesson, I intend to extract some further results that could be useful both in the case of a pure functioning of the market and in the case of well-meaning interventions.

The driving idea is that, perhaps, the structure of the network could be a powerful indicator of its potential performance. If that is the case, then the analysis of the network characteristics could serve a double goal. First, it would allow the authority to assess how inherently prepared (robust) that very structure is to withstand shocks, to defuse crises locally, and to avoid large welfare losses. Second, it could suggest the authority the set of incentives to promote or, vice versa, to prevent the emergence of specific structures. The ultimate explicit goal is the minimization of welfare losses

Fig. 7.3 Comparison of *laissez faire* and managed case

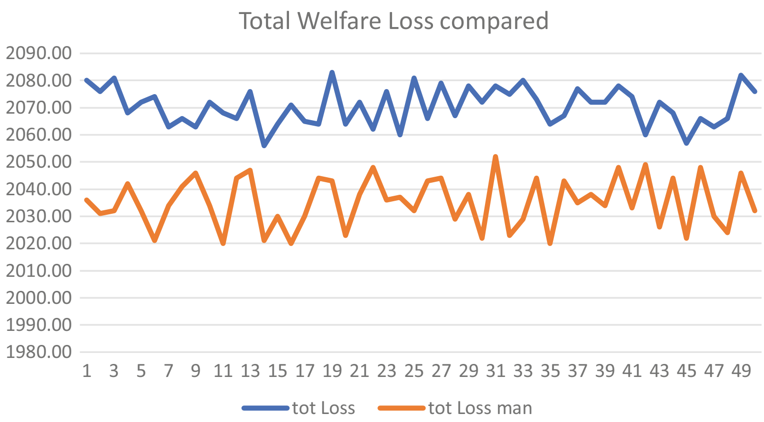


Fig. 7.4 Total social welfare loss

imposed by financial shocks. However, as shown above, the correct behavior is debatable and competing "intermediate" goals could result in structures that are "riskier": they could impose higher costs to the community.

Clearly, to be able to assess the potential performance of a network and/or to improve upon the current state, it is necessary to identify the characteristics of the "champions". That is, the structural features that have been consistently exhibited by the best performers and that can, hence, safely be employed as benchmarks.

It may seem that the champions could be identified by considering all possible structures. Well, that is certainly feasible for rather small networks. As an example, I have constructed—and tested against the same set of shocks—all possible networks of size 4, that is, with 4 nodes. A simple calculation shows that a directed graph with 4 nodes can be constructed in 2^{12} possible ways (4096).[5]

The equivalent calculation for the "toy model" of 15 nodes I use in the rest of the paper gives 2^{210} or 1,645,504,557,321,206,042,154,969,182,557,350,504,982,735,865,633,579,863, 348,609,024. It is clear that a brute force solution to the optimization problem for an actual financial system is not a simple task. However, 15 may not be such a ridiculously small number. For example, in Rwanda there are 16 commercial banks and a number of microfinance institutions and rural savings and credit cooperatives.

Back to the 4 node networks, I exposed each of them to the same set of 100 shocks and Fig. 7.5 represents the distribution of results in terms of average welfare loss.

The loss distribution in the two cases is quite interesting. In particular, the case of the managed allocation of shocks leads not only to a higher average for occurrence, but to a very volatile distribution. Furthermore, there is only one champion for the unmanaged case. It averages a loss of slightly over 3.5 units per crisis (which occurs on average every 2.37 "days"). On the other hand, there are 64 configurations that,

[5] As mentioned above, links from and to the same node are excluded so there are 12 possible connections. Each of these can be in two directions, hence the total 212

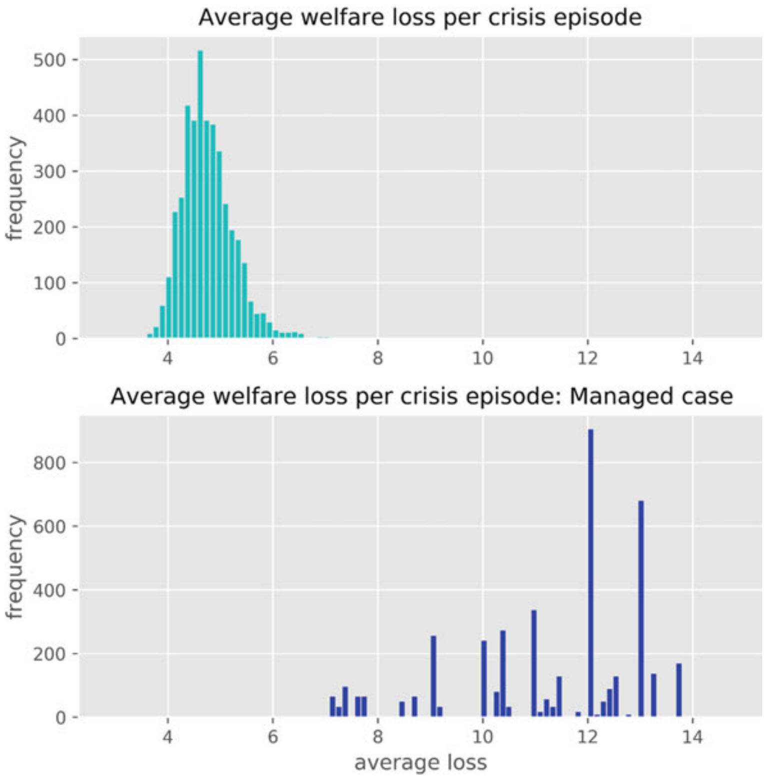


Fig. 7.5 Average welfare loss. Four-node networks

under the managed shock strategy all lead to the minimum average loss of 7.1 units per crisis (on average after 4.84 “days” in all cases).

A question arises when thinking of much higher network dimensions: how rugged is the performance landscape? How difficult would it be to identify (perhaps converge to) the optimum? Fig. 7.6 suggests that the landscape is highly nonlinear. Each network, represented by one point on the x axis, differs from its immediate neighbors only for one link.

Clearly the landscape is very rugged which implies that the optimization process (the search for the champions) is likely to get stuck in some local optimum, especially in the “managed” case where we see several large plateaus.

I approach the optimization problem in the following way. I create a population of networks and expose them to the same set of shocks. At the end of this set, I identify the network that has had the best overall performance and generate a new population in two different ways. The first one employs a rather crude genetic algorithm (it exploits only cross over), the second one, instead, applies simple random changes to

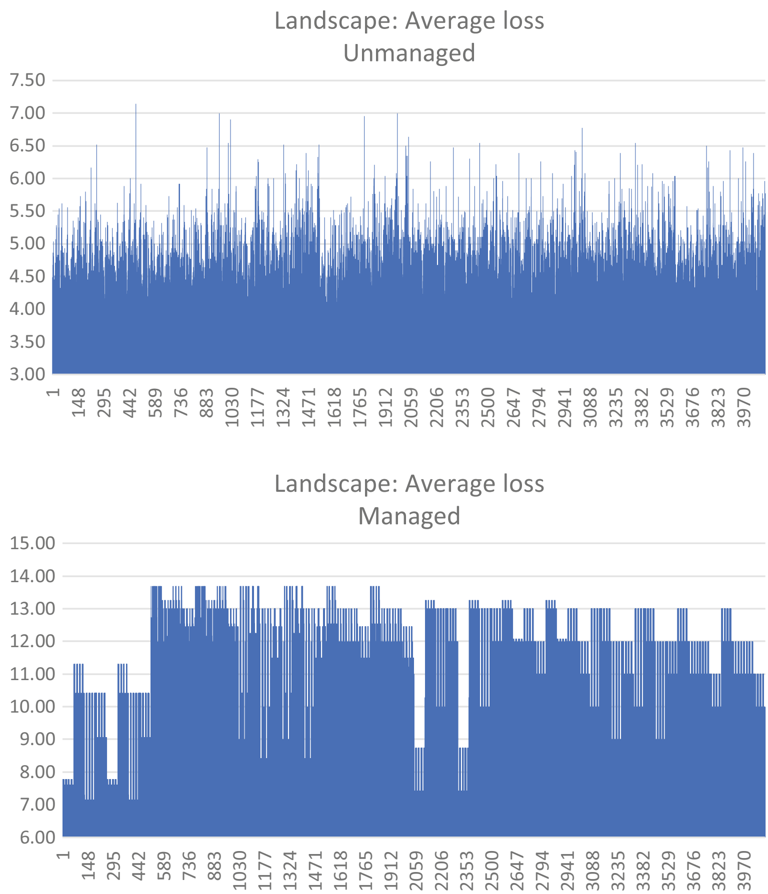


Fig. 7.6 Network landscape, 4 nodes

the original population (this could be interpreted as a rough mutation process). Both methods implement a search routine over a highly non-linear landscape.[6]

Each generation—composed of 26 "individuals"—is exposed to the same set of shocks and the selection process is repeated 30 times (each producing a new generation). At the end of the overall exercise the characteristics of the best network of the first generation are compared with those of the best network of the last generation

[6] If I had Professor Shu's expertise, I could have set up a two-stage approach. The first would have been based on genetic programming to search for an effective combination of routines (i.e. the genetic algorithm I employ or the random exploration, or the "gradient" approach sketched above). The result would have been a routine to apply in the selection process of the "best" network. So, I see genetic programming as a way to select algorithms (using the genetic principle of the survival of the fittest), while a genetic algorithm is only one such an approach to select the solution. I am aware that the distinction is more semantic than substantial, as a referee pointed out. However, it requires a programming competence that I do not own.

that, in my set up, has necessarily exhibited an equal or better overall performance. I say "necessarily" because I am using one of the tricks employed in these sort of simulations: my generations maintain a "ratchet"; that is, the best performer of the current generation remains the incumbent champion in the following generations until a better performer emerges. Furthermore, to make sure the results do not depend on the specific set of shocks employed, I also repeat the full set of exercises with a different series of shocks.[7]

The simulations with the same set of shocks and the same initial population are repeated under two different scenarios. In the first one, the shocks are assigned randomly to each node, while in the second case, the number of shocks per period is the same, but each new shock is assigned to one of the nodes with the least stock of NPLs.

It is important to note that the number of networks tested in this exercise, a total of 780 (26×30), is only a minuscule portion of the whole universe of networks. This is clearly a limitation on the possibility of identifying the absolute optimum, but it should still give an indication of the characteristics that would accompany the better performers.

The following series of pictures (Fig. 7.7) captures the starting and end value for several variables: (i) average time to crisis, (ii) average number of "slides" per crisis, (iii) average loss per crisis, (iv) total loss for the period considered, plus some structural features of the related networks: (v) size of the network, (vi) average clustering coefficient, and (vii) average minimum path of the network. All these data organized in a comparison between the "non managed" and "managed" case, using the GA or simple random search algorithms. There is also the special case called "moral hazard", that I discuss in the next section.

This set of results seems to confirm that there is a direct relation between average time to crisis and the magnitude of the average crisis event. However, they also seem to indicate that for a rather long period considered, the average time to crisis and total loss incurred are inversely related: the more frequent the crises, the larger the total loss. As for the characteristics of the network that could help infer the frequency and magnitude of the crisis episodes, it looks like the more "connected" the network, the less frequent but deeper the crises.

To support this anecdotical evidence, I have run a longer and more focused series of simulations. Considering only the genetic algorithm optimization method, I have performed three sets of 1000 simulations one each for the unmanaged, the managed, and the "moral hazard" case. Each one of the 1000 simulations starts with a different population and runs over 5 generations. The data of the champions at the end of each simulation give me 1000 observations for each approach. Figure 7.8 depicts the total loss distribution for each exercise.

It is clear that the "unmanaged" case—the pure *laissez faire* case—leads to the worst total loss performance. The managed case distribution does not even overlap

[7] I am not testing the champion emerged from one set of shocks on another set because my intent here is not the search for some "global" optimum, but rather try to identify the typical characteristics that a good network should have.

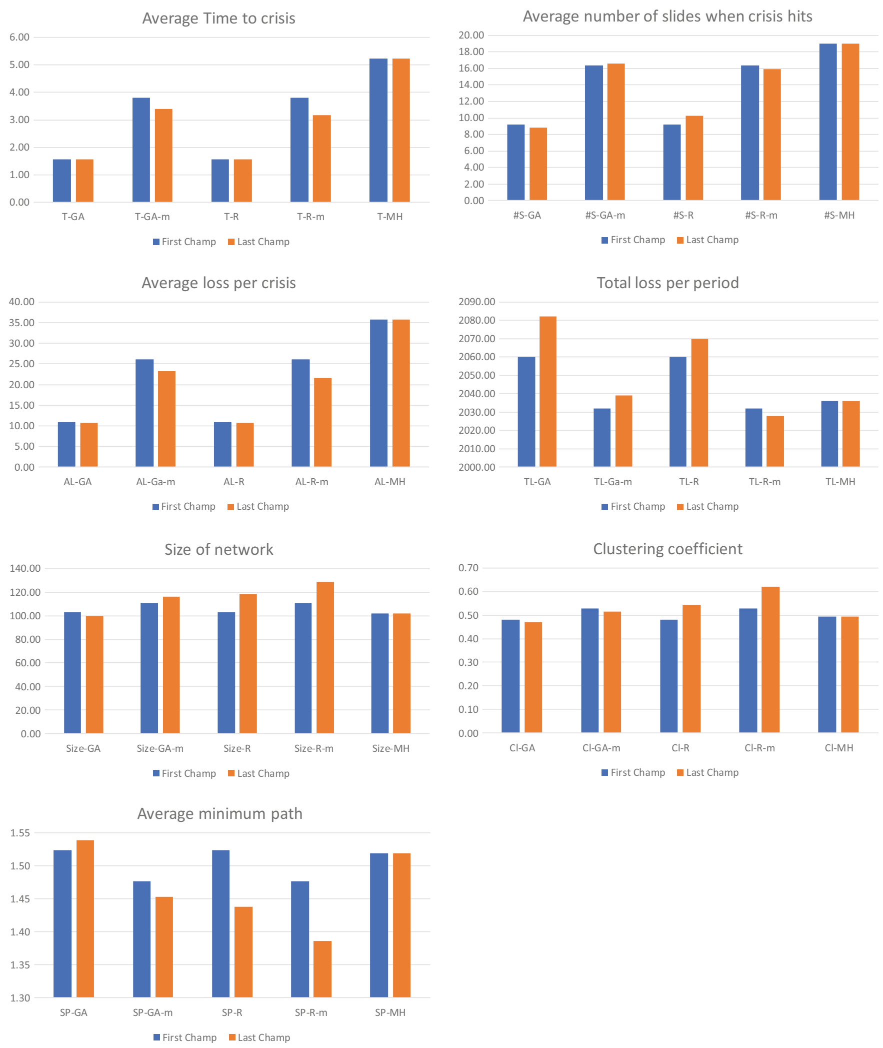


Fig. 7.7 Comparison of pre and post optimization variables

with the unmanaged one indicating an obvious improvement. Quite interestingly, the Moral Hazard case stands somewhat in the middle. Its best performance at 2020 is the same as the best performance recorded in the managed case and the frequency of such a performance is even higher. On the negative side, the volatility of the results is larger with the upper tail overlapping with the unmanaged scenario's results. However, considering that such a strategy is clearly the simplest to implement (once the authority has decided to get involved), the results are significant.

Moving now to the analysis of the characteristics of the networks in relation with total welfare loss, we start by providing a few simple correlations coefficients.

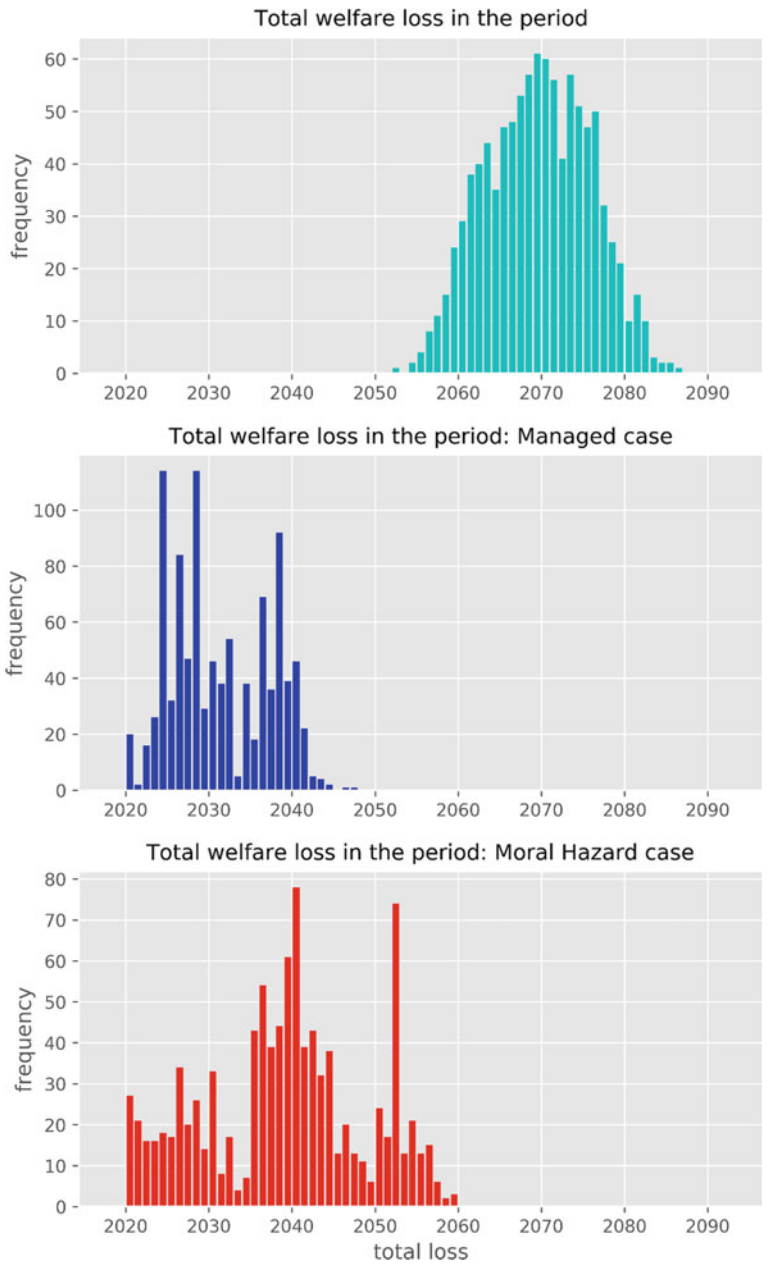


Fig. 7.8 Total loss distribution

Correlation between		
Average time to crisis and total loss		
Unmanaged scenario	Managed scenario	Moral hazard
–0.08	**0.41**	**0.18**
Network size and total loss		
Unmanaged scenario	Managed scenario	Moral hazard
–0.01	**–0.01**	**0.02**
Average cluster coefficient and total loss		
Unmanaged scenario	Managed scenario	Moral hazard
–0.006	**0.005**	**0.009**
Average shortest path and total loss		
Unmanaged scenario	Managed scenario	Moral hazard
–0.006	**0.015**	**–0.013**

The chosen network characteristics offered only some weak qualitative indications of the desirability of particular structures: the magnitude of the correlation coefficients is in all cases quite small. The only exception is the correlation of the average time to crisis to total loss in the managed case scenario. There seems to be quite a strong direct relation: less frequent crises, not only will be more costly on average, but also will quite often lead to relatively larger total welfare losses in the period considered. However, the relation I had initially foreseen between size and total loss and network connectedness (as represented by the cluster coefficient and the average shortest path length) and total loss appears negligible after these simulations.

Next, I performed two distinct econometric exercises, both employing OLS. In the first one total loss was the dependent variable, while in the second one I picked average loss.

In the first case, total loss was regressed on size of the network, average cluster coefficient, and average minimum path length. The same regression was repeated for each of the regimes (laissez faire, managed, and moral hazard). Unfortunately, the results (in Annex 1) were not conclusive at all. The R squared for all regressions is extremely low and the only variable that is consistently significant at the 5% confidence level is the intercept!

For the second[8] experiment—with average loss as dependent variable—I pooled the data across the three "regimes" and ran instead three equations, one for each independent variable: average size, average cluster coefficient, and average minimum path. The results (in Annex 2) add an interesting twist to the analysis. The original purpose was to identify the characteristics that the network should have to guarantee the best results in terms of minimum welfare loss in case of a crisis. The results of this exercise suggest that the winning structure depends on the prevailing regime.

[8] I thank my colleague Carlos de Resende for suggesting this approach, but do not hold him responsible for how I implemented it.

Not only the magnitude, but even the sign of the estimated parameter changes. So, for example, size plays a positive role in reducing the average loss in the laissez faire regime, but will increase the average loss in the moral hazard case.

7.6 Conclusions and Further Research Plans

The more general result obtained in this exercise is that the goal of welfare optimization unencumbered by other objectives, will lead the system to settle on a structure where crisis events are more frequent with a smaller average loss. However, "tampering" with the system so as to delay a crisis, leads to rarer, but, on average, more costly events. Surprisingly, the total cost (the sum of all costs caused by crises) during longer periods is, however, smaller for artificially obtained rarer events. The result would seem to suggest that the authority's direct intervention on the normal functioning of the system would have beneficial effects on welfare in the long run.

It is noteworthy that the magnitude of each rarer event is more uncertain. Hence, it requires larger provisions or more sophisticated insurance or hedging strategies. The longer intervals between crises, may also give the wrong incentive to myopic policy makers and lead to perverse behavior when combined with the election cycle.

I also considered an extreme case, which I denominate of *Moral Hazard*. In this scenario, the optimization criterion was the maximization of the time between crisis events or, equivalently, the minimization of the number of crises per period of time. The ensuing increase in the average magnitude of each (rarer) event was not surprising, but what was unexpected is that the total loss imposed on the public was not the largest of all cases considered and that the total loss distribution was in large part better than in the pure *Laissez Faire* scenario.

The attempt at inferring the potential performance of a network based on its inherent structural characteristics has led to some other unexpected results. The first one is that among the network characteristics I considered (size, clustering, minimum path) there does not seem to be one architecture that will always guarantee the best performance. Perhaps, the reason is that I have not yet identified this universally optimal feature. Or, as suggested in the second round of econometric analysis, the reason could be that the winning structure depends on the prevailing regime, so that features like size—that leads to a better performance in the *laissez faire* scenario—may be counterproductive in other regimes.

In any case, the analysis and results have suggested a series of questions and several possible extensions. The first one, in search for the best universal structure, is to apply the same methodology to a network that is structurally similar to a real financial system. This clearly implies to build not only a directed graph, but also a weighted one to capture more precisely the ownership and control relations among financial institutions. Furthermore, as mentioned, there are many other ways to characterize a graph that may be relevant to our exercise.

Annex 1

1. Unmanaged case, OLS: Total loss on Network size, Average cluster coefficient, and Average Shortest Path.

Summary output						
Regression statistics						
Multiple R	0.043607747					
R Square	0.001901636					
Adjusted R Square	−0.001104685					
Standard Error	6.280905725					
Observations	1000					
ANOVA						
	df	SS	*MS*	*F*	*Significance F*	
Regression	3	74.86138181	24.95379394	0.632545885	0.594103519	
Residual	996	39,291.97762	39.44977673			
Total	999	39,366.839				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	2103.071652	24.82160287	84.72747159	0	2054.363014	2151.78029
Size	−0.110851752	0.091947078	−1.205603862	0.228256602	−0.291283975	0.06958047
Cluster coefficient	2.014241673	10.70149477	0.188220591	0.850742072	−18.98582194	23.01430528
Average min path	−15.39680928	11.4159339	−1.348712197	0.177735998	−37.79885153	7.005232969

2. Managed case, OLS: Total loss on Network size, Average cluster coefficient, and Average Shortest Path.

Summary output						
Regression statistics						
Multiple R	0.043879967					
R Square	0.001925451					
Adjusted R Square	−0.001080797					
Standard Error	5.922861738					
Observations	1000					

(continued)

(continued)

ANOVA						
	df	SS	*MS*	*F*	*Significance F*	
Regression	3	67.40500171	22.4683339	0.640483108	0.589057049	
Residual	996	34,939.97	35.08029116			
Total	999	35,007.375				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	2027.507027	37.62180762	53.89180252	1.704E-297	1953.679924	2101.33413
Size	−0.074638441	0.125666276	−0.593941699	0.552685976	−0.321239485	0.171962603
Cluster coefficient	15.72815153	12.27210206	1.2816183783	0.200274822	−8.353991143	39.81029421
Average min path	2.289041938	17.99396398	0.127211655	0.898798558	−33.02138864	37.59947251

3. Moral Hazard case, OLS: Total loss on Network size, Average cluster coefficient, and Average Shortest Path

Summary output						
Regression statistics						
Multiple R	0.030871957					
R Square	0.000953078					
Adjusted R Square	−0.0020561					
Standard Error	9.706432875					
Observations	1000					
ANOVA						
	df	SS	*MS*	*F*	*Significance F*	
Regression	3	89.52020912	29.84006971	0.316723671	0.81330085	
Residual	996	93,837.97979	94.21483915			
Total	999	93,927.5				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	2000.366836	56.51876937	35.392996517	2.5369E-178	1889.457306	2111.276365
Size	0.172853045	0.197381122	0.875732406	0.381386667	−0.21447753	0.56018362
Cluster coefficient	−9.719035892	17.20984962	−0.564736828	0.57237994	−43.49076073	24.05268894
Average min path	16.55980766	26.09108755	0.63469212	0.525774994	−34.64000226	67.75961759

Annex 2

Dependent Variable: AVELOSS				
Method: Least Squares				
Date: 08/04/21 Time: 16:20				
Sample: 1 3000				
Included observations: 3000				
Variable	Coefficient	Std. Error	t-Statistic	Prob
C	23.3846	0.362415	64.52434	0
D1*SIZE	−0.121024	0.003463	−34.94762	0
D2*SIZE	0.000234	0.003372	0.06947	0.9446
D3*SIZE	0.13069	0.003526	37.0697	0
R-squared	0.982774	Mean dependent var		23.6553
Adjusted R-squared	0.982757	S.D. dependent var		10.80421
S.E. of regression	1.418729	Akaike info criterion		3.538732
Sum squared resid	6030.323	Schwarz criterion		3.546741
Log likelihood	−5304.098	Hannan-Quinn criter		3.541613
F-statistic	56,976.59	Durbin-Watson stat		2.012221
Prob (F-statistic)	0.000000			

Dependent Variable: AVELOSS				
Method: Least Squares				
Date: 08/04/21 Time: 16:27				
Sample: 1 3000				
Included observations: 3000				
Variable	Coefficient	Std. Error	t-Statistic	Prob
C	23.66775	0.338299	69.96106	0.0000
D1*AVECLUSTER	−25.76301	0.674256	−38.20955	0.0000
D2*AVECLUSTER	−0.495191	0.663853	−0.745935	0.4558
D3*AVECLUSTER	26.73404	0.688638	38.82163	0.0000
R-squared	0.981835	Mean dependent var		23.65530
Adjusted R-squared	0.981817	S.D. dependent var		10.80421
S.E. of regression	1.456899	Akaike info criterion		3.591830
Sum squared resid	6359.172	Schwarz criterion		3.599838
Log likelihood	−5383.745	Hannan-Quinn criter		3.594710
F-statistic	53,978.54	Durbin-Watson stat		2.014832
Prob(F-statistic)	0.000000			

Dependent Variable: AVELOSS				
Method: Least Squares				
Date: 08/04/21 Time: 16:30				
Sample: 1 3000				
Included observations: 3000				
Variable	Coefficient	Std. Error	t-Statistic	Prob
C	20.25031	0.715718	28.29369	0
D1*AVEMINPATH	−6.341986	0.472061	−13.43468	0
D2*AVEMINPATH	2.124291	0.479343	4.431671	0
D3*AVEMINPATH	10.92319	0.469567	23.26228	0
R-squared	0.987362	Mean dependent var		23.6553
Adjusted R-squared	0.98735	S.D. dependent var		10.80421
S.E. of regression	1.215182	Akaike info criterion		3.228997
Sum squared resid	4424.096	Schwarz criterion		3.237006
Log likelihood	−4839.496	Hannan-Quinn criter		3.231878
F-statistic	78,025.27	Durbin-Watson stat		2.052057
Prob(F-statistic)	0.000000			

References

Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: an explanation of $1/f$ noise. Phys Rev Lett 59(4):381–384
Bjorner A, Lovasz L, Shor PW (1991) Chip-firing games on graphs. Eur J Combin
Holroyd A, Levine L, Mészáros K, Peres Y, Propp J, Wilson B (2008) Chip-firing and rotor-routing on directed graphs. In out Equilib 2(60):331–364
Spencer J (1986) Balancing vectors in the max norm. Combinatorica 6:55–66

Chapter 8
A Systematic Review of Investor Attention: Measurements, Implications, and Future Directions

Dehua Shen and Chen Wang

Abstract Investor attention refers to the limited attention that investors can devote to the information which might affect their investment decisions. Given attention is a scarce cognitive resource and individuals have limited capacity to process information, existing studies have shown that the allocation of attention, which is caused by the attention-grabbing events, has impacts on financial market dynamics accordingly. These attention-grabbing (or "stimuli") events have also been adopted as indirect proxies to measure investor attention in early studies. This study contributes to the literature on investor attention in explaining why and how investor attention matters, especially to investor behaviours and financial market dynamics. It also reviews the measurements of investor attention used in current literature as well as the corresponding limitations. Based on a systematic review of the existing literature, this paper envisages potential future directions for investor attention with a more comprehensive understanding of current studies.

Keywords Investor attention · Market dynamics · Rational inattention · Market anomalies

8.1 Introduction

Attention is a scarce cognitive resource given the limited capacity in human cognitive systems, and it is a selective process in which the organism controls for the choice of stimuli and then individual behaviour (Egeth and Kahneman 1973). A host of psychological research studies human attention on handling multiple perceptual tasks simultaneously, and many theorists establish the basic concepts that the central

D. Shen (✉)
School of Finance, Nankai University, Tianjin, P.R. China
e-mail: dhs@nankai.edu.cn

C. Wang
College of Management and Economics, Tianjin University, Nankai District, No. 92 Weijin Road, Tianjin 300072, P.R. China
e-mail: saphirachen@tju.edu.cn

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_8

cognitive-processing capacity of human brains has limits (Pashler and Johnston 1998). Incorporating the realistic constraint that individuals' information-processing capacity is a limited resource, Sims (2003) builds a theoretical model of "rational inattention". He applies the concept of channel capacity from information theory to study information-processing constraints in a dynamic programming problem. In this setting, attention allocation or information processing has an opportunity cost, and the rational inattention theory becomes the intuition for studies afterwards.

It is therefore considered that attention to one task is a substitution for the attention to other tasks. This does not conflict with casual situations that people can watch a talk show while running on a treadmill, for instance, or simply talk over the smartphones while smoking and walking on the way back home. The difference here is between the simply perceptual analysis and the more central cognitive analysis involving memory retrieval and action planning required in multiple task operations (Peng and Xiong 2006). Based on this insight of mental operations, psychologists describe human brains as single-processor computers with switching and buffering capabilities, and the efficiency of generating an output depends on the processing time of the task which the computer allocated to (Pashler and Johnston 1998).

On the other hand, motivated by this psychological evidence, there are extensive studies examine the features of stimuli which attract investor attention (Odean 1998, 1999; Barber and Odean 2008 etc.). There is a tremendous amount of information involved in our daily lives, whereas salient information is able to attract more attention. In natural environments, salient stimuli are introduced as the salience stands out from its backgrounds or environments which could attract more attention and provide more influence on behaviour (Shinoda et al. 2001).

While in the context of financial markets, investors need to be selective both spatially and temporally in processing ever-growing amounts of information and make financial decisions. One of the first theoretical models for investor attention is advocated by Merton (1987) who incorporates the investor recognition hypothesis, assuming investors are only aware of a subset of available stocks in informationally incomplete markets. By analyzing market equilibrium and asset prices in this setting, he finds that stock price sometimes reacts to a widespread report containing all the substantive information which has already been announced previously. Typically, Huberman and Regev (2001) observe a representative example that a re-reporting article on the front page with "A special report" label on *New York Times* has made EntreMed (a small biotechnology company with the rights to commercialize in cancer-curing)'s stock price skyrocketed from 12 to as high as 85 before ending at around 52 on the first trading day after the release of the re-reporting article. However, this potential breakthrough in cancer research in *Nature* has already been reported by several popular newspapers (including *New York Times* over five months earlier). Adopting similar intuition, considering there are thousands of stocks in the market that are available for stock choices. Earlier scholars propose that it is attention, which takes the portfolio choice-set from several thousands of stock alternatives to a manageable number, and the final selection depends on investors' preferences and beliefs within this range of choices (Odean 1999).

Given investor attention a limited cognitive resource, a large number of studies have provided ample evidence showing that investor attention not only plays a crucial role in affecting investment decisions and other financial activities, but is also important in understanding market dynamics and predicting companies' future performance. Thus, the main objective of this study is to review the existing literature on investor attention in explaining why and how investor attention matters, and how to measure attention. To answer these questions, this study is structured as follows (see Fig. 8.1): as shown in Fig. 8.1, firstly, we begin by identifying attention-grabbing events (or the features of "salience") which could provide the first insight to investor attention. These events are consistent with indirect measures for attention in the next section. Secondly, we review the current indirect and direct proxies to measure investor attention, and the corresponding limitations in Sect. 8.3. Adopting these measurements for attention, the effects of investor attention/inattention on investor behaviour as well as financial market dynamics are examined in Sect. 8.4 to answer why and how attention matters. Then in Sect. 8.5, we provide some suggestions for future research in investor attention and conclude in the last section.

8.2 Attention-Grabbing Events

Attention-grabbing events, such as significant growth in trading volume and extreme abnormal returns, refer to the events or shocks which could attract investor attention within the market. Given the limited attention investors can devote into investments, and their likelihood to invest in stocks which catch their attention, researchers have studied extensively on the attention-grabbing events, or the features of salient stimuli in different circumstances.

In general, scholars try to observe the features of these "salience" and study investor behavior with different types of attention-grabbing events. Typically, attention-grabbing events can be classified as abnormal trading volume (Gervais et al. 2001; Barber and Odean 2008), extreme return (Koster et al. 2006; Seasholes and Wu 2007; Barber and Odean 2008; Yuan 2015; Gödker and Lukas 2019), advertising (Grullon et al. 2004; Lou 2014; Madson and Niessner 2019; Focke et al. 2020), news and headlines (i.e. media coverage) (Barber and Loeffler 1993; Liang 1999; Barber and Odean 2008; Engelberg and Parsons 2011; Yuan 2015), etc. We will review these events separately in the following paragraphs.

Firstly, abnormal trading volume is considered attention-grabbing. Since the trading volume of a stock largely depends on the number of traders, it is almost tautological that when more investors are actively trading a stock, more investors are paying attention to it (Barber and Odean 2008). According to Gervais et al. (2001), the visibility of stocks due to high trading volume would affect the demand and price of the stocks.

Secondly, previous studies document abnormal return as one of the key events which could attract investor attention. The intuition is that when there is a significant move in price, it is likely that no matter what leads to price move also catch investors'

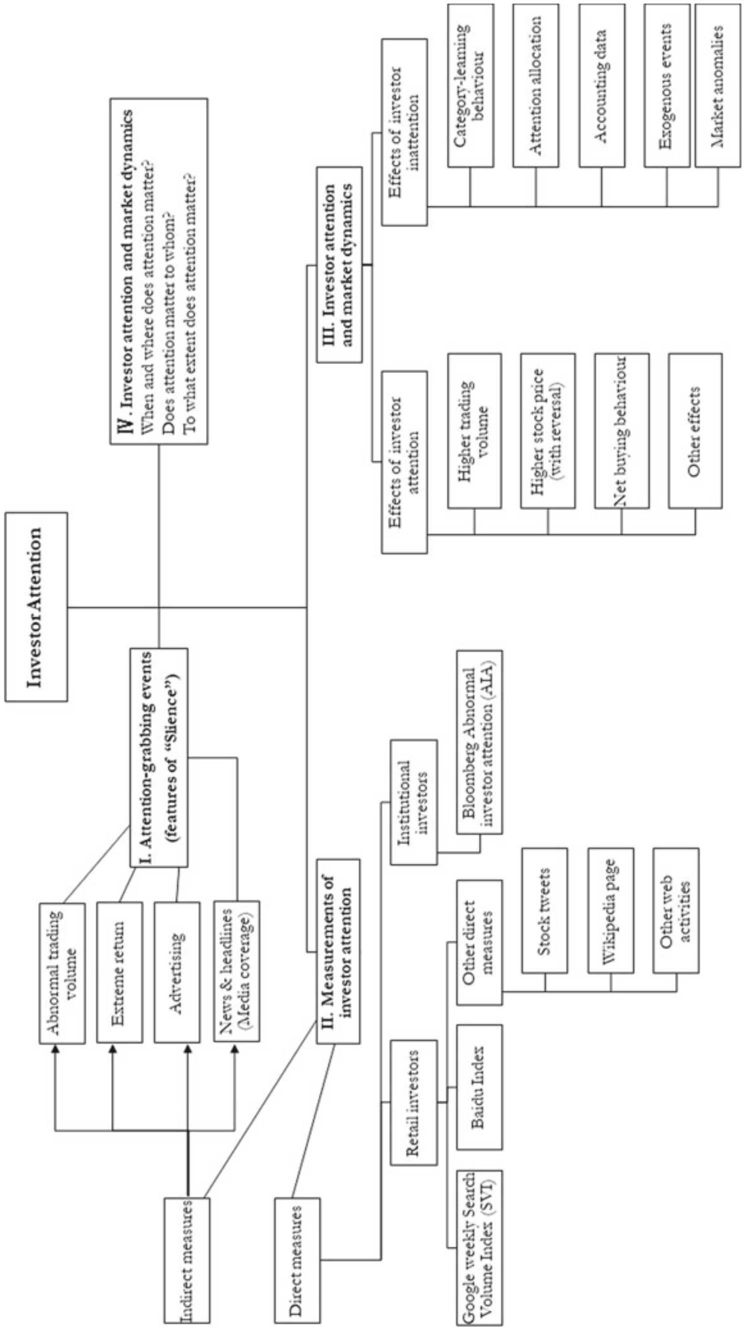


Fig. 8.1 Logic diagram of investor attention

attention (Barber and Odean 2008). To be specific, Seasholes and Wu (2007) observe the stock prices' hitting of upper price limit in Shanghai Stock Exchange as an attention-induced event, and make individual investors become net buyers in the stock market. On the contrary, Yuan (2015) examines the record-breaking events of Dow index and Dow history high indicator, and find these events would attract market attention, whereas individual investors tend to trade more actively but sell their holding stocks dramatically and lower the market price levels. In line with Barber and Odean (2008)'s propositions, Koester et al. (2016) observe that when companies (especially the ones with optical future performances) desire for attention, or if the managers consider their stocks are neglected by market participants, extreme positive earnings surprise could be a successful method to attract investor attention with a significant increase in trading volume, the number of institutional investors, and analysts in the following three years. A recent study from Gödker and Lukas (2019) uses an incentivized eye-tracking laboratory experiment, as the asymmetric mechanisms are associated with people's visual attention patterns. Rather than extreme positive return, their experiment indicates that individual investors are more likely to purchase stocks with recent extreme negative returns.

Another documented type of attention-induced events refers to advertising activities. At this point, Grullon et al. (2004) provide empirical evidence that companies with greater advertising expenditure can improve the overall visibility with investors, and therefore generate a larger number of both individual and institutional investors, and an increase in stock liquidity. Similarly, Lou (2014) analyzes the effect of companies' annual advertising spending on investor attention and finds that companies ranking top of advertising expenditures outperform those in the bottom line within three years. More recently, Madson and Niessner (2019) use companies' daily advertising data to examine the spillover effects of advertising on investors' attention in financial markets. Specifically, they find that print ads in business publications can trigger temporary spikes in investor attention, documenting that even frequent and repeated events which reveal little nonpublic information can also become investor attentive. Focke et al (2020) also find advertising positively affects investor attention, but has little impact on turnover and liquidity. Another study employs the corporate sponsorships of football bowl games as a natural experiment and finds that the sponsors' stocks have attracted more investor attention (an increase in Google search volume of those stocks) (Mayer 2020).

In regard to media coverage, with the same intuition as Barber and Odean (2008) who propose that companies covered in news and headlines are more likely to catch investors' attention than those that are not, it has been widely studied that media coverage has a causal impact in financial markets (eg., Fang and Peress 2009, Engelberg and Parsons 2011; Madsen and Niessner 2019). Also, Yuan (2015) suggests that unlike professional investors, a large portion of individual investors are relatively less sophisticated, and they are considered relying more on media coverage because they have limited accessibilities to several information channels. They find stocks covered in the front-page articles could capture more attention and generate higher market returns in the following day. Another strand of studies focuses on expert stock recommendations reported in the media. Although these recommendations rarely convey fresh information, they can sometimes play a role in bringing stocks into public sights. For example, Barber and Loeffler (1993) find that stocks recommended from 1988 to 2000 in "dartboard column" of the *Wall Street Journal* garnered attention and experienced a double in trading volume on the two days following the publishment. Liang (1999) also suggests a sharp increase in trading volume by examining the stocks which dartboard expert recommended from 1990 to 1994.

8.3 Measurements of Investor Attention

The main objective of this section is to review the common measurements (both indirect and direct) for investor attention. Measuring attention is a significant challenge since the inception of investigating investor attention. Some pioneer researchers resort to attention proxies as indirect measures which are closely related to investor attention, i.e., abnormal return, trading volume, media coverage and advertising. More recently, scholars also use direct measures such as Google search volume (SVI), Baidu Index, tweets and other direct proxies.

8.3.1 Indirect Measures

Asking hundreds and thousands of individual investors which stock or information they thought about or pay attention to in each and every past single day is considered an impossible task. As reviewed in Sect. 8.2, sufficient literature documents that investor attention is driven by attention-grabbing events such as excessive trading volume, abnormal returns, media coverage etc. Some of these events could both induce and follow excessive investor attention, and therefore become indirect proxies to measure investor attention.

Firstly, extreme returns are considered as an indirect proxy for investor attention. Barber and Odean (2008) are one of the pioneers to study the observable measures of investor attention. With the belief that investors are likely to be aware of the stocks which have extreme one-day returns, they document extreme return as one of the three "imperfect but useful" proxies for investor attention. In order to measure attention, Barber and Odean (2008) sort the stocks' daily return in the prior day into 10 deciles, and define extreme returns as the first decile (lowest returns) and last decile (highest return). In addition, Seasholes and Wu (2007) find a high proportion of purchases of stocks which hit upper price limit in Shanghai Stock Exchange are first-time buyers whose attention has been attracted by this event. Therefore, consistent with Barber and Odean's (2008) theory, they also propose that the attention-grabbing events are the "salience" among all the information available in the market which could attract and become indirect proxies for investor attention.

The second indirect proxy refers to trading volumes. Trading volume should be highly correlated with investor attention as active trading involves investors' attention to process information and analyze asset fundamentals (Hou et al. 2009; Gervais et al. 2001). Barber and Odean (2008) believe that trading volume is an indicator of attention which the stock is received. They propose abnormal trading volume as an indirect proxy for investor attention and find it the best measure of attention among the three indirect proxies they examined. In that sense, the abnormal trading volume is defined as the dollar volume for the stock traded on the day as reported in daily stock return files of CRSP for NYSE, ASE and NASDAQ, divided by the average trading volume. Similarly, Hou et al. (2009) adopt trading volume, measured as the average monthly turnover over (i.e., number of shares traded in a mont/number of shares outstanding at the end of month) the previous year, as a proxy for attention in the cross-sectional analysis to test the impacts of investor attention on stock price dynamics. Loh (2010) also uses the stock turnover as a proxy to investor attention in studying investor inattention.

Another strand of studies examines the relationship between investor attention and media coverage (news and headlines), and finds news and headlines as an indirect proxy for attention. Still, in order to test news coverage as an indirect proxy, Barber and Odean (2008) consider news associated with attention-grabbing events, and find that news is a primary mechanism for catching investors' attention. Specifically, they use the daily news feed from Dow Jones News Service containing the ticker symbols of the firms in the news, and sort the dataset into the ones with news covered and the ones without mentioned in daily frequency. There are limitations to this proxy as it is difficult to measure e intensity or salience of news coverage accurately (Barber and Odean 2008). As mentioned before, Yuan (2015) adopt front-page news event, which is defined as the situation when front-page stories about the stock market movements show on both *New York Times* and *Los Angeles Times*, as one of the attention-grabbing events which could also measure investor attention. In order to analyze the effects of the location of a lead IPO underwriter in investment banking networks, Bajo et al. (2016) use a hand-collected data set of pre-IPO media coverage including headlines and articles to measure investor attention.

In terms of advertising as an indirect proxy, besides attracting customers' attention and promoting sales of products, advertising can also direct investors' attention to the firm and trigger more possible trading of its stocks. Investors with higher attention to a firm are more likely to purchase its stocks. Based on this perception, scholars use advertising expenditure as the indirect proxy investor attention and indicate that a higher level of investor attention brings more following purchase decisions of the firms' stocks (Grullon et al. 2004; Chemmanur and Yan 2019; Lou 2014) and greater abnormal stock returns (Lou 2014). It is also documented that advertising in sports games as recognizable (Fehle et al. 2005) or sponsored firms (Mayer 2020), which attracts more attention, generates significant positive abnormal return following the game.

There are intrinsic limitations for the indirect measures reviewed above, as these proxies are primarily based on critical assumptions. For instance, as pointed out by Da et al. (2011), researchers often build theoretical models assuming that investors should have paid attention to the abnormal turnover, extreme return, advertising or media coverage of the stocks. However, investor attention also affected by other factors unrelated to investor attention such as macroeconomic environment (Engle and Rangel 2008), media coverage sometimes contains problems in independence (You et al. 2018) and local bias (Gurun and Butler 2012), and the release of news itself cannot guarantee the readiness by investors (see also Huberman and Regev (2001)'s vivid example).

8.3.2 Direct Measures

Comparing with indirect proxies with ex-post information, it is relatively difficult to adopt direct measures of investor attention in empirical studies, however, the advances in information technology and growing popularity of social media provide more opportunities for constructing more ex-ante proxies of measuring investor attention.

When investors intend to search for useful information relevant to a stock, they are largely interested or paying attention to that stock (Da et al. 2011; Chen, 2017, etc.). In a relatively timely fashion, Da et al. (2011) introduce Google weekly Search Volume Index (SVI) as a direct measure of investor attention, and define the abnormal Search Volume Index (ASVI) of a stock as the (log) SVI during the current week minus the (log) median SVI during the previous eight weeks. They observe that SVI, capturing less sophisticated individual investors' attention, is correlated with but different from existing measures of investor attention. Da et al. (2011) also find empirical evidence to support Barber and Odean's (2008) price pressure hypothesis that an increase in ASVI indicates a price increase in the following two weeks and an eventual price reversal within a year. ASVI is also positively correlated with the large first-day returns and long-run underperformance of IPO stocks.

Zhang et al. (2013) employ the search frequency of stock names in Baidu Index as the direct proxy for investor attention in China. Vlastakis and Markellos (2012) suggest SVI as a proxy for attention since it is also a proxy variable of information demand, and they observe a significant positive relationship between SVI and the past volatility and turnover of stock. After that, a strand of studies in the literature employs Google search data (or Google Trends) and Baidu Index to measure investor attention, and examine the dependence between attention and other financial information such as market volatility (Goddard et al. 2015; Dimpfl and Jank 2016), net-selling behavior (Chu et al. 2022), liquidity (Ding and Hou 2015) and returns of stocks (Da et al. 2011; Shen et al. 2017; Swamy and Dharani 2019; Li et al. 2021, etc.).

Other searching activities related to individuals' financial decisions have also been observed as direct proxies. For instance, Joseph et al. (2011) adopt online ticker searches to forecast abnormal stock returns and trading volumes, whereas the abnormal return and trading volumes are also indirect proxies for investor attention. Also, Mondria et al. (2010) use a new internet search query dataset from America Online (AOL) including over 21 million web searches by 657,426 customers as a direct measure for investor attention allocation, and they find empirical evidence of the two-way causality between investors' attention allocation and home bias.

There are also other direct measurements adopted by existing literature. More microscopically, Li et al. (2016) use asset-specific stock tweets containing more temporal information as a direct measure of investor attention, and sample from actual investors who regularly participate in the stock market. They quantify the "social power" of investor attention and find strong attention contagion effects in the current information environment. Moreover, Gargano and Rossi (2018) use investor web activities including the time spent on the research page of the brokerage account website, the number of web pages browsed and the number of logins to individual accounts to measure investor attention. In a similar vein, Focke et al. (2020) propose another measure of attention based on the daily number of page views of the firm's Wikipedia page, considering its advantage in unambiguous identification of the firm, which could largely distinguish investor attention from consumer attention; and also advantage in easier comparison across firms and time. As for the cryptocurrency market, Shen et al. (2019) employ the number of tweets from Twitter as a direct measurement of attention of uninformed investor and find that the number of tweets is a significant driver of next day realized volatility and trading volume. More recently, Iliev et al. (2021) use investors' views of proxy filings in EDGAR to measure investors' attention to corporate governance, which could isolate attention from other financial and operating issues.

On the other hand, there are also measures of institutional investors. For instance, having considered that direct investor attention measures such as google search activity and stock tweets capture mostly retail investor attention, Ben-Raphael et al. (2017) propose a novel direct proxy for institutional investor attention using daily search and readiness of news for specific stocks on Bloomberg. As there are almost 80% users of Bloomberg terminals work in financial industries, they define abnormal institutional attention (AIA) as a dummy variable to reveal institutional investor attention. The construct of AIA makes it easier to compare institutional and individual investor attention in different settings. For example, Ben-Raphael et al. (2017) show that institutional attention responds more quickly to major news events and leads to more trading activities than individual investor attention. They also find that institutional attention facilitates permanent price adjustment, which is opposite to the price pressure effect (i.e., increase and following reversal in price) of retail investor attention to stock price (Barber and Odean 2008; Da et al. 2011 etc.).

8.4 Investor Attention and Market Dynamics

Based on the efficient market hypothesis (EMH), stock price should respond instantaneously to new information and therefore affect investment characteristics. However, in practice, absolute market efficiency can be hardly achieved because investors are unable to process every piece of information in a timely manner. Important information is not fully incorporated into stock price until investors pay attention to it (Merton 1987; Huberman and Regev 2001). The attention-grabbing information can also improve market valuations by alleviating informational frictions (Merton 1987).

Theoretical and empirical studies suggest that attention is one of the key determinants of the statics and dynamics of asset returns (Da et al. 2011; Andrei and Hasler 2015). There is a host of studies suggesting that how individuals allocate attention affects their investment behaviours and therefore has a series of economic consequences.

8.4.1 Effects of Investor Attention

There is ample evidence showing that investor attention plays a crucial role in affecting investor behaviour and therefore has accordingly impacts on financial market dynamics. As we will review in this section, the gathering of investor attention would generate more trading volume (Mayer 2020), realized volatility (Zhang et al. 2021a), higher stock price (with reversal) (eg., Barber and Loeffler 1993; Seasholes and Wu 2007; Chen 2017; Gödker and Lukas 2019; Mayer 2020; Zhang et al. 2021b etc.), net buying behaviour (Odean 1999; Barber and Odean 2000, 2008), etc.

Firstly, it is reasonably considered that investors cannot actively trade in the stock market with no attention paid to it. As reviewed in last section, abnormal trading volume has been defined as one of the attention-grabbing events. Conversely, when investor attention is higher, the abnormal trading volume would be more pronounced (Mayer 2020).

Investor attention can also affect stock returns. As we have already discussed, stocks with extreme returns are relatively easier to attract investor attention no matter such returns are positive or negative. However, the literature documents that attention-driven purchases of stocks by individual investors often inflate stock price temporarily, but lead to lower subsequent returns or price reversals. There is a strand of studies examining the effect of investor attention generated by experts' recommendations on the price movements of stocks. Early in 1993, Barber and Loeffler find that the attention-driven purchases of stocks recommended by experts on the *Wall Street Journal* leads to an average positive abnormal return of 4 percent, and the price of these stocks partially reversed within 25 trading days. Wright (1994) also report similar findings with a 4.59% abnormal return in two days and a complete price reversal after 36 trading days. Moreover, Liang (1999) documents a 3.5% abnormal return within 2 days' announcement window and a price reversal in 15 days. As one of the hot papers in this field, Barber and Odean (2008) propose the attention-induced price pressure hypothesis arguing that the net buying of attention-grabbing stocks can lead to over-reaction to information. Da et al. (2011) support this argument with the evidence that an increase in search frequency in Google (SVI) predicts higher stock prices in the two weeks but an eventual price reversal within the year. In a similar vein, Seasholes and Wu (2007) find individual investors in Shanghai Stock Exchange who buy stocks following attention-grabbing events lose 0.88% due to the price mean-revert over the next five days, whereas the stock prices embrace a transitory impact but reverse to pre-event levels within ten trading days. Quite the contrary, in a broader context, Chen (2017) explores the effects of investor attention on global stock returns, and finds an index returns decrease significantly following the increase of investor attention, and such negative effects of investor attention would be enhanced in the market with positive sentiments. Gargano and Rossi (2018) also find that attention is positively related to the future performance of the stocks (up to 4 months).

Thirdly, the gathering of investor attention can lead to the buy-sell imbalance. There are thousands of alternatives available in the stock market. Unlike professionals, most individual investors face search problem in stock selection. They have less time and resources to monitor and evaluate a wide range of stocks. Based on this concept, behavioral theories indicate that attention-grabbing events help to narrow the range of individual stock selections to a "consideration set" (Odean 1999). As most individual investors devote only limited attention with holding of three or four stocks to their portfolio. This makes the effects of investor attention weaker in selling decisions as the number of stocks holding in the portfolio are manageable and individual investors rarely sell short (Barber and Odean 2000). Besides, Barber and Odean (2008) proposed the asymmetrical effect of investor attention on buying and selling adopting the same basics in their previous studies. They suggest that although fully rational investors or professionals could realize that the attention-grabbing events have already been incorporated into the current stock price, which might have no relevance with future performance, there is still a notion of general attention effect in financial markets that (individual) investors would like to purchase stocks which have recently captured their attention. Thus, individual investors become net buyers of attention-grabbing stocks. Also, as mentioned previously, Seaholes and Wu (2007) show in their study that higher investor attention could not only make active individual investors become net buyers of the stocks, but also lead those investors who have never owned a stock in the past to purchase under this circumstance. However, the attention-based buying disappears when many events happen simultaneously where the search costs for the stocks are not reduced. The causal link between investor attention and stock purchase behaviour at the individual level has also been examined and proved by Gödker and Lukas (2019), Mayer (2020), etc. On the other hand, the institutional investors, especially value managers exhibit an opposite tendency of the individual investors, whereas they purchase more on days of low abnormal trading volume than high-volume days, and they are net buyers of stocks following daily extreme returns (Barber and Odean 2008).

Scholars also observe other kinds of effects generated from investor attention. For instance, Li et al. (2016) identify the social proof effect of investor attention in the modern information environment as the drawing power of the crowd make investor attention no longer isolated. Moreover, Hsieh et al. (2020) examine the effect of attention on the herding behaviour, and observe the asymmetric impact of investor attention since individual investors' herding behaviour in purchase becomes stronger in bull markets, especially for stocks with small market capitalization. Bajo et al. (2016) studies the IPO underwriter networks and consider more investor attention a significant channel for more central lead IPO underwriters to achieve a series of favorable IPO characteristics such as larger initial returns, greater institutional investments and stock liquidities, etc. Interestingly, Gargano and Rossi (2018) employ a novel brokerage account data set and construct various measure of attention. They find more attentive investors achieve better portfolio performance with greater risk-adjusted returns and Sharpe ratios. More recently, rather than focusing on capital markets, Kupfer and Schmidt (2021) study the effect of investor attention on odd lot trading. They observe that the relationship between attention and odd lot trading is price-sensitive: odd lot trading is negatively related to investor attention when stock price is less than $11, while they move in the same direction when stock price is over $46.

8.4.2 The Implications of Investor Inattention

While there is ample literature studying the attention-grabbing events for market participants, a strand of literature focuses on investor inattention, which also subjects to limited investor attention and cognitive resources, and information processing abilities. On one hand, investors may devote too much attention to some of the information which results in overreaction effects to the market, on the other hand, investors may also direct too little attention to other significant information and lead to delayed (or under-) reactions.

8.4.2.1 Effects of Investor Inattention

Rational inattention theory indicates that investors tend to allocate more attention to market-level shocks than firm-specific shocks, resulting in high stock return co-movement. Consistent with this concept, Peng and Xiong (2006) provide a theoretical model which shows that the limited investor attention leads to investors' category-learning behaviour, which is generalized as investors would allocate more attention to broader information (i.e., market and sector information) rather than firm-specific information. To further test this theoretical prediction and examine the determinants of investor attention, Liu and Peng (2015) compare investor attention to earnings announcements on days with and without macro news announcements, and find consistent evidence that investors would prioritize their attention to market-wide information when it comes simultaneously with firm-specific information shocks. Moreover, when multiple earnings announcements are issued on the same day, investors could be distracted with lower attention paid to each announcement. In sum, they observe that investors could allocate their limited attention strategically to process information and make investment decisions accordingly. Similarly, Drake et al. (2016) introduce the concept of attention comovement in their study, and document that the attention paid to a specific firm is correlated to the attention paid to its industry and the market in general. This also generates market consequences as the investor attention comovement is positively associated with the comovement of excess stock return.

In terms of attention allocation for portfolio management, Peng (2005) observes that individual agents in her continuous-time equilibrium model tend to allocate more attention and time to larger stocks in their portfolio, as these stocks contribute greater uncertainty to the portfolio. This also indicates that the price of large stocks tends to incorporate fundamental shocks faster than the small ones. As for attention allocation to different stocks within the market, Chakrabarty and Moulton (2012) examine the attention constraints of a market marker by analyzing the effect of one firm's earnings announcement on the liquidity of other firms' stocks including the ones in different industries. Consistent with their conjectures, the non-announcement stocks have significantly lower levels of liquidity on days when other stocks release earnings announcements. Other literature examines attention allocation in other contexts such as mutual funds (Kacperczyk et al. 2016; Gupta-Mukherjee and Pareek 2020).

There is sufficient evidence on the limited attention to accounting data. Many studies in this strand show that investor attention could mislead investors in reaction to accounting information. For example, Hirshleifer and Teoh (2003) make early effort in examining the limited attention effects on some publicly available accounting information which the firms can choose alternative methods to present, such as financial disclosure, financial reporting policy, and they analyze the effects of different presentations on market trading and prices as well. Still motivated by investors' limited attention as a cognitive constrained resource, Hirshleifer et al. (2004) suggests that investors tend to overvalue the firm with "bloated" balance sheet highlighting positive aspects of firms' performance and find the accounting ratio that net operating assets scaled by total assets can be a strong negative predictor of long-term stock return. Moreover, Hirshleifer et al. (2010) further their research with a focus on accounting information. In another theoretical model based on investors' inattention to current earnings, different earnings components and other investment related information, they find a stronger underreaction to cash flow than to earnings, and an overreaction to accruals when earnings information is attentive. It also shows that in market equilibrium, some investors may choose not to be attentive to some of the accounting information owing to the cost of time and attention. From another perspective, Drake et al. (2012) study the effect of investor attention to earnings announcements and find that when more attention is allocated to the firm in advance, there is less response to the announcement in price and volume, suggesting the earlier investor attention partially preempts the information content of earnings announcement. Focusing on corporate governance, Iliev et al. (2021) also observe investors tend to devote substantial attention to a certain subset of firms, so that the others might receive relatively little external monitoring.

Another strand of studies explores exogenous events which could distract investors from financial markets to examine the implications of investor attention. For example, Huang et al. (2019) find evidence from repeated natural experiments in Taiwanese stock market using large jackpot lotteries as exogenous shocks of investor attention that the rational inattention is extant, as the stock return co-movement with the entire market goes up on large jackpot days. Hu et al. (2021) observe opposite findings. From another perspective, Israeli et al. (2021) use the daily news pressure index (DNP) capturing breaking news as an indicator to unexpected events which distract investor attention, and find investors are susceptible to distractions in their reactions to earnings announcements especially for retail investors.

8.4.2.2 Investor Inattention as a Source of Market Anomalies

Scholars tend to use investor (in)attention to explain several market anomalies in financial fields such as earnings momentum (Hirshleifer et al. 2009; Dellavigna and Pollet 2009; Hou et al. 2009; Dellavigna and Pollet 2009) and price momentum (Hou et al. 2009).

Investor inattention can help to explain investors' underreaction to news and announcements, whereas less attention results in a delayed response, but a stronger post-earnings announcement drift (Hirshleifer et al. 2009; Dellavigna and Pollet 2009). Typically, when firms are in a competitive environment for investor attention, Hirshleifer et al. (2009) build a model to show that the limited attention leads to investors' mis-reaction to the earnings surprises and the accrual anomaly, while the initial price and volume reactions are weaker and post-earnings announcement drift much stronger when other firms release earnings announcements in the same day. Similarly, Dellavigna and Pollet (2009) find supportive evidence for explaining the presence of post-earnings announcement drift by the underreaction to information due to limited investor attention and inattention hypothesis. The intuition here is that investors tend to be distracted by the upcoming weekend and pay less attention to investment-related information. They compare investors' reactions to earnings announcements on Fridays and on other weekdays, and show a 15% lower immediate response and a 70% higher delayed response. Investors would eventually observe the mispricing problem and incorporate Friday information into prices. In addition, still using Friday as an exogenous variation in the level of attention, Louis and Sun (2010) study a type of unscheduled attention-grabbing events and observe investors' underreaction to merger announcements made on Fridays, which is in line with the hypothesis proposed by Dellagigna and Pollet (2009).

Hou et al. (2009) also examine the effects of investor attention on asset price dynamics through investors' under- and overreactions to information using trading volume as the proxy for attention. They consider that investors may ignore useful information which leads to price underreaction, and therefore provide a possible explanation for the post-earnings announcement drift (i.e., earnings momentum) as the ignored information gradually gets incorporated into prices and drift in the same direction with the earnings announcement news. On the other hand, Hou et al. (2009) assume investors may also be attentive to stocks and their related information which leads to stock price overreaction to generate price momentum. They find profits are higher for price momentum strategies among high volume stocks, while earnings momentum profits are higher for low volume stocks.

Besides assuming investors' underreaction to news and announcements due to their less attentiveness, investor inattention (or limited attention) is also used to explain investors' underreaction in other settings. For instance, Loh (2010) examines the effect of investor inattention on investors' response to stock recommendations using stock turnover as the proxy for investor attention. Having considered inattention as a plausible explanation for post recommendation drift, he finds that stocks with low investor attention react less to stock recommendations than those with high investor attention around a 3-day event window. On the contrary, Drake et al. (2012) consider this phenomenon caused by investors' abnormal attention to the earnings information beforehand, and thereby preempts the content of earnings announcements at the release day. More recently, Jiang et al. (2019) analyze the effect of investor attention on 17 widely studied anomalies in financial markets and show causal evidence that anomaly returns are higher during the days following the increase in investor attention, which also provides potential arbitrage opportunities.

8.5 Future Directions

The existing literature mainly focuses on "why" and "how" attention matters. However, future research could also look for the answers to "when and where does attention matter?", "does attention matter to whom?", or "to what extent does attention matter?" For example, understanding the fundamental logic and features of investor attention would be helpful for researchers to examine it in different contexts. Existing studies have started to observe the characteristics of investor attention. Specifically, supporting the assumptions about investor attention made by Dellavigna and Pollet (2009) and Hong and Yu (2009), Liu and Peng (2015) demonstrate strong seasonality of investor attention. They examine the patterns of attention directly and find investor attention significantly lower on Fridays and in summer months (July and August). In addition, Chen (2017) finds the geographic patterns of investor attention. He suggests an asymmetric attentional allocation to market by local investors versus foreign investors due to the local information advantages. Besides the characteristics of investor attention such as seasonality, geographic patterns and herding behaviours (Hsieh et al. 2020), researchers may also look for other features of investor attention, and other potential applications of these features in investors' financial behaviours.

As for the "to whom" question, the majority of current studies focus on individual investors, whereas much fewer studies examine institutional investor attention (Barber and Odean 2008; Ben-Raphael et al. 2017). A few scholars have started to examine institutional attention recently in several contexts such as different clientele effects in investor attention (Liu et al., 2021), the effects of institutional investor attention on the timeliness of analysts' forecasts for the firm (Chiu et al. 2021) and on managers' disclosure adjustments (Abramova et al., 2020), etc. As institutional investors actually play a vital role in financial markets (Ben-Rephael et al. 2017, 2021), although still hard to quantify, it is worth studying for researchers in comparison to retail investor attention.

In terms of "to what extent", a great number of studies are based on critical assumptions that investors should have paid attention to attention-grabbing events, or the shocks of financial market generated from more attention could also become proxies to measure investor attention. In extent to Da et al. (2011) and Ben-Raphael et al. (2017), the advance in information technology and machine learning could provide more opportunities for future researchers to study or examine investor attention in a novel and more accurate way. For instance, Iliev et al. (2021)'s novel empirical measure of attention on corporate governance could avoid the uncertain elements within the general measures of attention. Future studies could also construct other measures of attention for more specific issues and contract attention by different types of investors.

Furthermore, given limited investor attention, besides the existing literature on the effects of exogenous events on attention allocation (Huang et al. 2019; Hu et al. 2021; Israeli et al. 2021), how to allocate attention efficiently and effectively in practice, and the corresponding asset pricing implications would also be interesting for further research. On the other hand, besides considering from investors' perspective, attention may also be studied from the managerial or public firm side, or it can be observed from analysts' responsiveness as well in regard to investor attention.

8.6 Conclusion

As documented by sufficient literature, attention is a scarce cognitive resource and individuals' capabilities of processing information is limited (Chen 2012; Tai et al. 2018). This chapter has reviewed the investor attention related literature in several sectors such as the possible attention-grabbing events (i.e., the features of "salience") including extreme stock return, abnormal trading volume, advertising and news/media coverage etc. Measuring attention is a significant challenge for researchers referring to its inherent characteristics. Earlier scholars tend to use excessive trading volume, abnormal return, media coverage and advertising as indirect measures of investor attention. More recently, with the advances in information technology and growing popularity of social media, scholars introduce direct measures of investor attention such as Google search volume, Baidu Index, stock tweets, search and readiness of news on Bloomberg etc. Moreover, as attention affects investment behaviours such as significant change in trading volume and buy-sell imbalance, and therefore it becomes one of the key determinants of the statics and dynamics of the financial market. The limited investor attention also leads to investor inattention, which could be used as explanations of several market anomalies such as stock price co-movements and post-earning announcement drift. This study makes a comprehensive review on the basics of investor attention, existing measurements of attention and effects of attention on market dynamics. It contributes to the literature in related to investor attention mainly answering "why" and "how" attention matters, and provides potential directions for future research in this field for scholars who are interested.

Acknowledgements This work is supported by the National Natural Science Foundation of China (72071141).

References

Abramova I, Core J, Sutherland A (2020) Institutional investor attention and firm disclosure. Account Rev 95(6):1–22

Andrei D, Hasler M (2015) Investor attention and stock market volatility. Rev Financ Stud 28(1):33–72. https://doi.org/10.1093/rfs/hhu059

Bajo E, Chemmanur TJ, Simonyan K, Tehranian H (2016) Underwriter networks, investor attention, and initial public offerings. J Financ Econ 122(2), 376–408. https://doi.org/10.1016/j.jfineco.2015.12.001

Barber BM, Loeffler D (1993) The "Dartboard" column : second-hand information and price pressure. J Financ Quant Anal 28(2):273–284

Barber BM, Odean T (2008) All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev Financ Stud 21(2):785–818. https://doi.org/10.1093/rfs/hhm079

Barber BM, Odean T (2000) Trading is hazardous to your wealth: the common stock investment performance of individual investors. J Financ 55(2):773–806. https://doi.org/10.1111/0022-1082.00226

Ben-Rephael A, Carlin BI, Da Z, Israelsen RD (2021) Information consumption and asset pricing. J Financ 76:357−394

Ben-Rephael A, Da Z, Israelsen RD (2017) It depends on where you search: institutional investor attention and underreaction to news. Rev Financ Stud 30(9):3009–3047. https://doi.org/10.1093/rfs/hhx031

Chakrabarty B, Moulton PC (2012) Earnings announcements and attention constraints: the role of market design. J Account Econ 53(3):612–634. https://doi.org/10.1016/j.jacceco.2012.01.001

Chemmanur TJ, Yan A (2019) Advertising, attention, and stock returns. Q J Financ 9(3). https://doi.org/10.1142/S2010139219500095

Chen S-H (2012) Varieties of agents in agent-based computational economics: a historical and an interdisciplinary perspective. J Econ Dyn Control 36(1):1–25. https://doi.org/10.1016/j.jedc.2011.09.003

Chen T (2017) Investor attention and global stock returns. J Behav Financ 18(3):358–372. https://doi.org/10.1080/15427560.2017.1331235

Chiu P, Lourie B, Nekrasov A, Teoh SH (2021) Cater to thy client: analyst responsiveness to institutional investor attention. Manag Sci. https://doi.org/10.1287/mnsc.2020.3836

Chu G, Li X, Zhang Y (2022) Information demand and net selling around earnings announcement. Res Int Bus Financ 59:101522. https://doi.org/10.1016/j.ribaf.2021.101522

Da Z, Engelberg J, Gao P (2011) In search of attention. J Financ 66(5):1461–1499. https://doi.org/10.1111/j.1540-6261.2011.01679

Dellavigna S, Pollet JM (2009) Investor inattention and friday earnings announcements. J Financ 64(2):709–749. https://doi.org/10.1111/j.1540-6261.2009.01447

Dimpfl T, Jank S (2016) Can internet search queries help to predict stock market volatility? Eur Financ Manag 22(2):171–192

Ding R, Hou W (2015) Retail investor attention and stock liquidity. J Int Finan Markets Inst Money 37:12–26

Drake MS, Jennings J, Roulstone DT, Thornock JR (2016) The comovement of investor attention. Manag Sci 63(9):2847–2867. https://doi.org/10.1287/mnsc.2016.2477

Drake MS, Roulstone DT, Thornock JR (2012) Investor information demand: evidence from google searches around earnings announcements. J Account Res 50(4):1001–1040. https://doi.org/10.1111/j.1475-679X.2012.00443

Egeth H, Kahneman D (1973) Attention and effort. Am J Psychol 88(2). https://doi.org/10.2307/1421603

Engelberg JE, Parsons CA (2011) The causal impact of media in financial markets. J Financ 66(1):67–97. https://doi.org/10.1111/j.1540-6261.2010.01626

Engle RF, Rangel JG (2008) The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. Rev Financ Stud 21(3):1187–1222. https://doi.org/10.1093/rfs/hhn004

Fang L, Peress J (2009) Media coverage and the cross-section of stock returns. J Financ 64(5)

Fehle F, Tsyplakov S, Zdorovtsov V (2005) Can companies influence investor behaviour through advertising? super bowl commercials and stock returns. Eur Financ Manag 11(5):625–647. https://doi.org/10.1111/j.1354-7798.2005.00301

Focke F, Ruenzi S, Ungeheuer M, Engelberg J, Harvey C, Odean T, Mai V, Schneemeier J, Schuster P (2020) Advertising, attention, and financial markets acknowledgements: we are grateful to. Rev Financ Stud. https://academic.oup.com/rfs/advance-article-abstract/. https://doi.org/10.1093/rfs/hhz142/5658636

Gargano A, Rossi AG (2018) Does it pay to pay attention? Rev Financ Stud 31(12):4595–4649. https://doi.org/10.1093/rfs/hhy050

Gervais S, Kaniel R, Mingelgrin DH (2001) The high-volume return premium. J Financ 56(3):877–919. https://doi.org/10.1111/0022-1082.00349

Goddard J, Kita A, Wang Q (2015) Investor attention and FX market volatility. J Int Finan Markets Inst Money 38:79–96

Godker and Lukas (2019) Attention to extreme returns. Working Paper.

Grullon G, Kanatas G, Weston JP (2004) Advertising, breadth of ownership, and liquidity. Rev Financ Stud 17(2):439–461. https://doi.org/10.1093/rfs/hhg039

Gupta-Mukherjee S, Pareek A (2020) Limited attention and portfolio choice: the impact of attention allocation on mutual fund performance. Financ Manag 1–43. https://doi.org/10.1111/fima.12294

Gurun UG, Butler AW (2012) Don't believe the hype: local media slant, local advertising, and firm value. J Financ 67(2):561–598. https://doi.org/10.1111/j.1540-6261.2012.01725

Hirshleifer DA, Lim SS, Teoh SH (2010) Limited investor attention and stock market misreactions to accounting information. SSRN Electron J. https://doi.org/10.2139/ssrn.846685

Hirshleifer D, Teoh SH (2003) Limited attention, information disclosure, and financial reporting. J Account Econ 36(1–3 SPEC. ISS.):337–386. https://doi.org/10.1016/j.jacceco.2003.10.002

Hirshleifer D, Hou K, Teoh SH, Zhang Y (2004) Do investors overvalue firms with bloated balance sheets? J Account Econ 38(1–3 SPEC. ISS.):297–331. https://doi.org/10.1016/j.jacceco.2004.10.002

Hirshleifer D, Lim SS, Teoh SH (2009) Driven to distraction: extraneous events and underreaction to earnings news. J Financ 64(5):2289–2325. https://doi.org/10.1111/j.1540-6261.2009.01501

Hong H, Yu J (2009) Gone fishin': seasonality in trading activity and asset prices. J Financ Markets 672–702

Hou K, Xiong W, Peng L (2009) A tale of two anomalies: the implications of investor attention for price and earnings momentum. Working Paper. https://doi.org/10.2139/ssrn.976394

Hsieh SF, Chan CY, Wang MC (2020) Retail investor attention and herding behavior. J Empir Financ 59(July):109–132. https://doi.org/10.1016/j.jempfin.2020.09.005

Hu Y, Li X, Goodell JW, Shen D (2021) Investor attention shocks and stock co-movement: substitution or reinforcement? Int Rev Financ Anal 73. https://doi.org/10.1016/j.irfa.2020.101617

Huang S, Huang Y, Lin TC (2019) Attention allocation and return co-movement: evidence from repeated natural experiments. J Financ Econ 132(2):369–383. https://doi.org/10.1016/j.jfineco.2018.10.006

Huberman G, Regev T (2001) Contagious speculation and a cure for cancer : a nonevent that made stock prices soar. J Financ 56(1):387–396

Iliev P, Kalodimos J, Lowry M (2021) Investors' attention to corporate governance. Rev Financ Stud. https://doi.org/10.1093/rfs/hhab003

Israeli D, Kasznik R, Sridharan SA (2021) Unexpected distractions and investor attention to corporate announcements. Rev Acc Stud. https://doi.org/10.1007/s11142-021-09618-4

Jiang L, Liu J, Peng L, Wang B (2019) Investor attention and asset pricing anomalies. SSRN Electron J. https://doi.org/10.2139/ssrn.3437527

Kacperczyk M, Van Nieuwerburgh S, Veldkamp L (2016) A rational theory of mutual funds' attention allocation. Econometrica 84(2):571–626

Koester A, Lundholm R, Soliman M (2016) Attracting attention in a limited attention world: exploring the causes and consequences of extreme positive earnings surprises. Manag Sci 62(10):2871–2896. https://doi.org/10.1287/mnsc.2015.2286
Kupfer A, Schmdit MG (2021) In search of retail investors: the effect of retail investor attention on odd lot trades. J Empir Financ 62(2021):315–326
Li X, Hendler JA, Teall JL (2016) Investor attention on the social web. J Behav Financ 17(1):45–59. https://doi.org/10.1080/15427560.2015.1095752
Li Y, Goodell JW, Shen D (2021) Comparing search-engine and social-media attentions in finance research: evidence from cryptocurrencies. Int Rev Econ Financ 75:723–746. https://doi.org/10.1016/j.iref.2021.05.003
Liang B (1999) Price pressure: evidence from the "Dartboard" column. J Bus 72(1):119–134. https://doi.org/10.1086/209604
Liu H, Peng L, Tang Y (2021) Retail attention, institutional attention. Working Paper
Loh RK (2010) Investor inattention and the underreaction to stock recommendations. Financ Manag 39(3):1223–1252. https://doi.org/10.1111/j.1755-053X.2010.01110
Lou D (2014) Attracting investor attention through advertising. Rev Financ Stud 27(6):1797–1829. https://doi.org/10.1093/rfs/hhu019
Louis H, Sun A (2010) Investor inattention and the market reaction to merger announcements. Manag Sci 56(10):1781–1793
Madsen J, Niessner M (2019) Is Investor attention for sale? The role of advertising in financial markets. J Account Res 57(3):763–795. https://doi.org/10.1111/1475-679X.12257
Mayer EJ (2020) Advertising, investor attention, and stock prices: evidence from a natural experiment. Financ Manag 1–34. (April 2018). https://doi.org/10.1111/fima.12324
Merton RC (1987) American finance association a simple model of capital market equilibrium with incomplete information. J Financ 42(3):483–510
Mondria J, Wu T, Zhang Y (2010) The determinants of international investment and attention allocation: using internet search query data. J Int Econ 82(1):85–95. https://doi.org/10.1016/j.jinteco.2010.04.007
Odean T (1998) Volume, volatility, price, and profit when all traders are above average. J Financ 53(6):1887–1934. https://doi.org/10.1111/0022-1082.00078
Odean T (1999) Do investors trade too much? Am Econ Rev 89(5):1279–1298. https://doi.org/10.2307/j.ctvcm4j8j.28
Pashler H, Johnston JC (1998) Attentional limitations in dual-task performance. January
Peng L (2005) Learning with information capacity constraints. J Financ Quant Anal 40(2):307–329. https://doi.org/10.1017/s0022109000002325
Peng L, Xiong W (2006) Investor attention, overconfidence and category learning. J Financ Econ 80(3):563–602. https://doi.org/10.1016/j.jfineco.2005.05.003
Seasholes MS, Wu G (2007) Predictable behavior, profits, and attention. J Empir Financ 14(5):590–610. https://doi.org/10.1016/j.jempfin.2007.03.002
Shen D, Urquhart A, Wang P (2019) Does twitter predict Bitcoin? Econ Lett 174:118–122. https://doi.org/10.1016/j.econlet.2018.11.007
Shen D, Zhang Y, Xiong X, Zhang W (2017) Baidu index and predictability of Chinese stock returns. Financ Innov 3(4):1–8. https://doi.org/10.1186/s40854-017-0053-1
Shinoda H, Hayhoe MM, Shrivastava A (2001) What controls attention in natural environments? Vision Res 41(25–26):3535–3545. https://doi.org/10.1016/S0042-6989(01)00199-7
Sims CA (2003) Implications of rational inattention. J Monet Econ 50(3):665–690. https://doi.org/10.1016/S0304-3932(03)00029-1
Swamy V, Dharani M (2019) Investor attention using the Google search volume index – impact on stock returns. Rev Behav Financ 11(1):56–70. https://doi-org.ezproxy.library.sydney.edu.au/. https://doi.org/10.1108/RBF-04-2018-0033
Tai C-C, Chen S-H, Yang L-X (2018) Cognitive ability and earnings performance: evidence from double auction market experiments. J Econ Dyn Control 91:409–440. https://doi.org/10.1016/j.jedc.2017.12.008

Vlastakis N, Markellos RN (2012) Information demand and stock market volatility. J Bank Financ 36:1808–1821
Wright DW (1994) Can prices be trusted? A test of the ability of experts to outperform or influence the market. J Acc Audit Financ 9(2):307–323. https://doi.org/10.1177/0148558X9400900211
You J, Zhang B, Zhang L (2018) Who captures the power of the pen? Rev Financ Stud 31(1):43–96. https://doi.org/10.1093/rfs/hhx055
Yuan Y (2015) Market-wide attention, trading, and stock returns. J Financ Econ 116(3):548–564. https://doi.org/10.1016/j.jfineco.2015.03.006
Zhang W, Shen D, Zhang Y, Xiong X (2013) Open source information, investor attention, and asset pricing. Econ Model 33:613–619. https://doi.org/10.1016/j.econmod.2013.03.018
Zhang W, Yan K, Shen D (2021a) Can the Baidu Index predict realized volatility in the Chinese stock market? Financ Innov 7(7):1–31. https://doi.org/10.1186/s40854-020-00216-y
Zhang Y, Chu G, Shen D (2021b) The role of investor attention in predicting stock prices: the long short-term memory networks perspective. Financ Res Lett 38:101484. https://doi.org/10.1016/j.frl.2020.101484

Chapter 9
What is the Market? The Essential Teachings from an AI Market Experiment

Yuji Aruka

Abstract So far traditional economics chose to argue "market in general", as symbolized by general equilibrium theory, which was originally established in 1870's. In reality, unfortunately, there is no longer a held "generality" by any means, as Leon Walras firstly expected. To depict the modern market, it may be not only inappropriate but also boldly far-reaching for us to discuss the market in general. Fortunately, we recently acquired two major technological advancements. One is AI in general, the other is Bitcoin or cryptocurrency in particular. The latter is a byproduct in association with blockchain technology. But the impact of Bitcoin is enormous in theory and in practice. This has the potential power to innovate the economic system, quite possibly, the social system. This article focuses on the settlement mechanism of the Bitcoin Exchange to discuss its uniqueness and commonality as compared with the stock exchange. We then introduce our new idea of digit length frequency distribution when we are interested in fully random iterated cellular automata (FRICA). Finally, we apply the length distribution to the price time series generated by the U-Mart acceleration experiment tool and examine the neutralizing effects by the change of the market transaction strategy composition, which may suggest some possible criteria whether the market may be rigged or not.

Prof. Chen is a distinguished professor of economics and complexity science. I am pleased to be given the honor to post my small submission to celebrate his 60th anniversary. He is internationally well-known Genetic Programming. There are shared research interests between the two of us. AI economics is one of Chen's main subject areas, both for education and research. I have been dealing with a similar subject, independently of Chen's AI-ECON research center (http://www.aiecon.org) and been committed to the U-Mart project, which was developed by Japan Association for Evolutionary Economics, both for education and research. I hope that this AI market experiment chapter would an appropriate tribute in his festschrift and expect to have new research conversations with Chen in the future.

Y. Aruka (✉)
Chuo University, Hachioji, Tokyo 192-0393, Japan
e-mail: aruka@tamacc.chuo-u.ac.jp
URL: https://yuji-aruka.jp

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_9

Keywords U-Mart system · Acceleration experiment · Digit length frequency distribution · Roll of non-intelligent agent · Neutralizing effect

9.1 The Limitations of the Libertarian Idea of the Market and Our New Interest

The libertarian idea, at a glance, seems attractive, especially during the numerous current contemporary change in the world. However, this idea, once we observed a vastly transformed infrastructure due to new scientific innovation, may be unfortunately unreliable. The classical market view does not hold in the new supply-demand relationship, which is easily understandable if we observe the customer service relationship of the IT industry. In this production process, the customers will be an indispensable input component. This is a main characteristics common for the service providing industry, as detailed in Aruka (2015). In this circumstance, the customer/consumer is never positioned in a one-way direction outside of production. Customers remain in the production process, instead of dropping unilaterally outside the production for pure consumption. Customers' capability and participation imply resources for producing the final output, whether physical or nonmaterial. Our social state has already transformed into a service-dominated society. Moreover, the ratio of service production to GDP amounts to 80 percent in many advanced countries.

This situation is not one supposed by the classical market view in which the independency of consumer preference from production is guaranteed. We do not state the modern situations around the new infrastructure more. Unfortunately, the libertarian subjectivism is usually deprived of its existence due to the economic system structure in itself, which excludes the independency of preference subjectivism. This structural change was not simply brought by the advent of an AI dominated society. We may not overlook the spread of national scaled innovation led by the syndicate of government-industry-university to enjoy the AI-dominated society.[1] Without such superforces to support the incessant promotion of innovations, AI infrastructure cannot be maintained/developed.

We can also reason another aspect of the limitation of preference subjectivism. Individualistic preference theory applies only to consumers who procure consumer goods by procuring them from producers, whether directly or indirectly, from retailers via wholesalers. However, over 70–80 percent of the national proceedings of the domestic production will be attributed to the transactions among producers[2] To start, traditional economics, so far, excessively emphasized individualistic choice as the role of coordination force of the market. In modern times, this role is moving from consumption to production. Furthermore, it is also important to point out that the

---

[1] It is easily understandable if we learn from the power gird development at the catch-up stage of the industry in the earlier stage of capitalism.

[2] In Japan, the consumption is approximately ¥300 trillion of the GDP. But the total sales of the domestic production is approximately ¥1000 trillion.

democratic society is changing its meaning. Even the procedural property, such as voting, depends on a voting machine algorithm in the era of AI identification. However, we do not refer to this issue any further. We simply point out the effect of propaganda power on the market and democracy. If some force such as propaganda power exists, then formal reasoning based on the game rules will then be deprived of generality and fairness.

In the early 20th century, the advent of the mass market, of necessity, induced propaganda into the mass society. Oxford Dictionary defines "propaganda" as information, especially of a biased or misleading nature, used to promote a political cause or point of view. In 1928, Edward Bernays[3] sharply pointed out : "The conscious and intelligent manipulation of the organized habits and opinions of the masses is an important element in democratic society." Bernays (1928, 9). Once this propaganda apparatus is set in the democratic process, propaganda will lead to the emergence of powerful elites in the modern bureaucratic system, capturing the decision-making cycles of our government, and rigging them towards their own private ends. Sometimes this will reinforce some social agreement promptly. This decision loop will also be entailed in the market process to direct a particular interest. There is no area that is exempt from any manipulation of the democratic or market process.

The outcomes that actually appear are always accompanied by a wide variety of different meanings despite referring to the same event. In this world, a formal outcome is just a "pie in the sky". The design decoration strategies on the mechanism provides a special benefit — room for procurement. The tax hole is a well-known example. The so-called "naked short hedge", often previously used on the Wall Street,[4] is also well-known . This is hinted at by skillfully rigging the holes of the stock investment institutional setting for the purpose of exchange. This kind of hole may also not be able to be fixed. The relevant legal regulation will be easily overturned as we have AI-related progress and other cunning knowledge.[5] Thus here we will not proceed to discuss some regulations to make up some system holes lawfully and economic theoretically.[6] In fact, there are a flourishing number of articles that discuss the regulation of cartel formation and the use of the AI development. However, the AI progress will not only invalidate these attempts (See Crandall et al. (2018)) but also provide a more profitable opportunity for professional traders. Thus, we will also not refer to this problem. More importantly, we should be liberated from the idealistic view on the market and its optimal achievement.

Thus, we will present actual market observations. More specifically, we confine our subject to a special interest in the settlement mechanism of the stock exchange and the effects of AI on the exchange market.

---

[3] Bernays was a "double nephew" of Sigmund Freud, a great psychoanalyst, on either maternal or paternal side. He is called "the father of public relations".

[4] See **naked short selling** in *Deep Capture* https://www.deepcapture.com/tag/naked-short-selling/.

[5] Remember a classical example of David Ricardo's observation of the quality of land. The quality of land , i.e., fertility, is changed by the progress of fertilizer.

[6] See Economic Justice and Natural Law (2009, 85–6), Crowe and Barbora (2016) and so on.

9.2 Settlement Mechanism of the Exchange Market

9.2.1 Preliminaries

Bitcoin is a very noteworthy entity in its property and role in the economic system (Nakamoto 2019). The largest reason why we should be interested in Bitcoin is that the trust is placed in computer technology and cryptography, not in humans, more specifically, the monetary authority or central banking system. However, the purpose of this article is not to argue the innovative properties of Bitcoin/Ethereum and other crypto-currencies and a theoretical issue such as **Byzantine general problem**. These have been already demonstrated in Aruka et al. (2020). In Aruka et al. (2020, Sect. 1), the differences were between the auction systems and the exchange market systems and between the market with auctioneer and without auctioneer.[7]

Bitcoin Exchange Bearing the reality of the exchange in mind, it may be better to take an example of the Bitcoin Exchange. **The Bitcoin Exchange** usually operates for 24/7, and there is no price movement limit.[8] Moreover, there are **big players** who can quickly manipulate the going rate through bidding.[9] It is noted that they are high-frequency traders with enormous amounts of money. Most noteworthy, this exchange market provides us with complete digital access. In this exchange, an average nonprofessional user who places his/her own order directly via the Internet not via an over-the- counter sale, can make transactions at any time.

In general, the settlement in the market is said to realize *equilibrium* at a price that suggests equality between the ask (supply) and the bid (demand). In the following, we apply **the English auction rule** of bidding from the bottom up, instead of the Dutch rule. There will be two main settlement mechanisms: batch and continuous double auctions. For simplicity, we take the double auction mechanism. We then have two methods (strategies) to ask and bid orders: a **limit order** and a**market order**. The former involves a combination of price and quantity, while the latter simply involves quantity at a going price. According to the order method used, the result will differ. Whether a limit order or a market order, any agent is allowed to place an order, either using an ask or a bid, as well as simultaneously using an ask and a bid.[10]

[7] Morishima (1984) is still an excellent textbook to carefully expose the theoretical settlement mechanism in the commodity exchange intermediated by auctioneers where sellers and buyers are equal partner. Furthermore, we have two kinds of markets: buyers select sellers, sellers select buyers. (Aruka et al. (2020, Sects. 1.2 to 1.3)).

[8] The price movement limit is an important restriction to suppress any sudden surge of price. This limit is usually adopted in the stock exchange. Without this limit, there is allowed to increase bigger. This is a reason why the Bitcoin price movement could be violent.

[9] Rubaiyat Islam in his Ph.D. dissertation (Islam 2021) challenged to de-anonymize the Bitcoin transaction data.

[10] Taken at 8:19am, March 3 2021 LTP: Latest Traded Price.

9.2.2 Numerical Examples at the Bitcoin Exchange

Market order We take an example from a current session from the Bitcoin Exchange of Bitflyer.[11] See Fig. 9.1. If he employs the limit order to ask at 0.5 unit with ¥5,146,694, he/she must wait for a buy order occurs that is at least greater than ¥5,146,694 with an order of greater than 0.5 units. Otherwise, he/she will be able to employ the market order of 1.08 units. If he puts his order of 0.5 units by **the market order** at this point in time, then the traded price will decrease to ¥5,142,212, because of the bidding column on the right hand side of Fig. 9.1: the combination of buy order sets (price and quantity) aligns from top to bottom: (5142303, 0.872), (5142301, 0.198), (514212, 0, 01). That is, just at ¥5,142,212, during the event, the demand quantity can amount to the supply order of 1.08 units. In our example, the latest traded price (LTP) is ¥5,144,931. Thus, this settlement will lead to a decreased equilibrium price. During the process attaining the final equilibrium of this case, it is noted that 0.872 units are first-matched at ¥5,142,303, and then 0.198 units are second-matched at ¥5,142,301. These prices are called partial equilibria, and these settlements are called piecewise stipulations. We remark that many piecewise stipulations occur in the market[12]

Limit order 1 If the agent bid his order of 0.5 units at ¥5,143,000, simultaneous another agent asks his order 0.2 units at the same price, it is partly stipulated at this price but with 0.3 units excess demand remaining. In this case, the piecewise stipulated equilibrium price is still below the latest traded price (LTP).

Limit order 2 If the agent follows **the limit order** to ask 0.5 units at ¥5,141,200, 0.01 unit is first matched at ¥5,141,212, s at ¥5,142,301 and third ¥5,142,303. Hence, the latter is the final equilibrium price, but with the excess demand of 0.362 units not being eliminated at this equilibrium price.

Limit order 3 We then deal with a large amount of orders. Suppose that an order will be placed for 100 units of the ask (sell) price prescribed at ¥5,000,000. Perhaps a bid order greater than ¥5,000,000 will be short of 100 units of selling. The sell order at this price will be forced to wait for the accumulated amount of the bid orders to be 100 in total. In this situation, a more bid orders at slightly greater than ¥5,000,000 may may be induced. It is noted that there are usually 1000 orders in per millisecond in digital transaction. Thus, within a second, this large deal will be settled, and then, the traded price will decrease to ¥5,000,000.

The effect of a large dealing In our custom of economic teaching, we are inclined to neglect a large dealing, as a large transaction will influence a change rate in the price. If there is a limit order of a large amount of ask, then it will imply a large decrease in the traded price, as mentioned above. However, the opposite does not necessarily hold. A large change in price with a small amount of bids, such as 0.05 units, but at ¥5,147,838, the settled price will be suddenly hiked up to ¥5,147,397. This hike may not be cleared because the ask corresponding to

[11] https://bitflyer.com/en-jp/.

[12] The settlement mechanism of piecewise stipulation in the case of double auction without the price movement limit will make the exchange a large commission rate for the exchange.

Fig. 9.1 Bitcoin transaction table

ask	price	bid
0.5	5 147 838	
0.02	5 147 397	
0.06	5 147 392	
0.5	5 146 694	
0.08	5 146 582	
0.131	5 145 671	
LTP	5 144 931	0.06%
	5 142 303	0.872
	5 142 301	0.198
	5 141 212	0.01
	5 141 111	0.01
	5 141 096	0.1

this price may still remain just 0.45 (=0.5 − 0.05). It is natural to observe that this traded price will be blocked if any bid at a greater price than this traded price is given in any amount. If this bid is slightly larger than 1 unit, then the final price must be decrease to ¥5,147,392. However, at this price, there is still an excess supply of 0.08 units.

Taking into account the piecewise stipulation (piecewise stipulated equilibrium) , it is seen that a higher bidding amount will not necessarily affect the final traded price. Therefore, we first learn that the existence of only a large dealing will not necessarily manipulate price movement. Second, the popular idea of the law of supply and demand will also not necessarily be held. Sometimes, a higher bidding will finally result in a lower traded price.[13]

9.2.3 A Time-Scale Problem

To date, there are often seen some controversies on how long a time window should be to understand its properties. However, it is interesting to learn the different time-scaled presentations of the same price movement. We show the graphs taken from Bitflyer's current transaction Figs. 9.2, 9.3, 9.4 and 9.5. It is seen from these figures that traders must be forced to live simultaneously in the different time scaled worlds

[13] We have the idea of **ask depth** and **bid depth** at any instant of transaction, which is the accumulated amount of orders at this instant. This will help as an assistant index to predict the price time series.

Fig. 9.2 Transaction graph aggregated every 1 min, March 5, 2021

Fig. 9.3 Transaction graph aggregated every 10 min, March 5, 2021

Fig. 9.4 Transaction graph aggregated every 1 hour, March 5, 2021

in order to predict the price movement. Chaotic movement will be analyzed in a later section of this article in terms of Turing machine.

Remark When an individual participates into the exchange, he/she will not usually feel the result from this settlement fair or good, except for he/she has procured positive earnings. Individuals learn to have a set of opportunities to either partly affect or manipulate the direction of price movement partly in partly his/her discretion if

Fig. 9.5 Transaction graph aggregated every 1 day, March 5, 2021

he/she has enough funds to invest in the exchange market.[14] Although its outcome has never been confirmed, the market is simply a machine for traders to generate any result under quickly changing circumstances and heterogeneous interacting agents with different expectations. This picture is quite different from the static situation preferred in economics.

9.3 AI Market Experiment

9.3.1 Experiment by the Use of the U-Mart System

As we have seen for the market exchange, like the Bitcoin exchange, it is difficult to make good or bad judgements about a result accruing from any market transaction. Surplus for both the bidder and provider is said to be maximized in equilibrium. However, the idea of surplus is a relative idea to be measured by the difference between a bid (or ask) price and the equilibrium price. In other words, surplus will be the difference between a piecewise stipulated price and the final stipulated price, as stated above. Assume that a preference list either for the bidder or for the provider is independently given. However, bidding prices are easily overturned instantly by a new swarm of bidders. It is difficult to see the independent supply and demand curves to continue to hold even in the short run.[15] We know that it is not possible to calculate economic surplus when we do not have the independent supply and demand curves. In this sense, in the actual market place, the meaning of optimality is

[14] GAMESTOP incident is an interesting event in the exchange market by See Wikipedia article "GameStop short squeeze" happened in January 2021:https://en.wikipedia.org/wiki/GameStop_short_squeeze.

[15] It is verified that the so-called double auction mechanism will not provide the information needed to construct the independent supply and the independent demand curves. It is easily understandable if you design AI exchange transactions such like U-Mart experiment. See Aruka (2015).

not clear. Thus, our market experiment will be irrelevant for verifying the optimality of market achievement.

In our view, the market is just a machine or server that provides individuals with a high-tech field (arena) to receive order inputs, match mutually, and send out outputs under a highly intensive high-frequency load. This process is no longer human-centrically observed. As emphasized in the above section, we do not have any automatic system such as an "invisible hand", either to achieve any kind of optimality or to certainly predict the future. In this section, thus, we execute the AI market experiment by using **the U-Mart system**, which is our long-run used market simulator.[16]

9.3.2 Nakajima-Mori Agent Configuration for the Generation of Price Movement Similar to Its Referential Price Time Series

We have already confirmed some remarkable results through a series of empirical method called **the acceleration experiment toolkit**. That is, Aruka et al. (2019, 2020) presented a special set called the **Nakajima-Mori agent set** to realize the future price time series similar to any given spot price time series in the AI market experiment. The Nakajima-Mori agent set fulfills the following two remarkable properties:

1. the existence of a minimum size for the agent configuration to realize a given price movement.
2. the scalar multiple of the agent size distribution will not change the result.

Now, we represent the Nakajima-Mori agent composition of any number, any scaled number and the mini number in Table 9.1.

The Nakajima-Mori agent set is empirically derived by detecting the stratum composition of a generated price movement which realizes the minimum of deviation from an externally given price time series. It is also proven in Aruka et al. (2021) that this composition can produce an approximate time series almost everywhere with a certain admissible range of errors. More interestingly, the minimum set will play the roll of the necessary condition to generate similar price movement if a certain strategy having an affinity with the minimum set is successfully added. We expect that some adaptive agent strategy will be a candidate.

[16] The U-Mart system is an artificial intelligent futures transaction system, which has emerged in the last 20 years, constructed by Japanese computer scientists and evolutionary economists. This system is compatible with batch and double auction mechanisms. See Shiozawa et al. (2008), Kita et al. (2016).

Table 9.1 The Nakajima-Mori agent composition of any number, any scaled number and the mini number

Agent strategy	Original	Any scaled	Min bodies
RandomStrategy	8.1	4	1
AntiTrendStrategy	7.2	3	1
DayTradeStrategy	8.2	4	1
MovingAverageStrategy	26.4	12	3
RsiStrategy	9.3	4	1
TrendStrategy	13.2	6	2
*S*RandomStrategy	50.8	23	6
*S*FSpreadStrategy	45	20	5
*S*MovingAverageStrategy	32.8	15	4
*S*RsiStrategy	22.6	10	3

As we have seen in Aruka et al. (2021), the **SFSpreadStrategy** will contribute to a similar movement, sometimes with a closer movement to the referential spot price time series. The experiment of the minimum set added to 20 bodies of the SFSpreadStrategy realizes almost the same as the referential spot price time series, sometimes closer to it than the series with the minimum set. Here, we also show that the **SFSpread strategy** will play the same role. In the present case, we prepared the case of the minimum set added to 10 bodies of the SFSpreadStrategy. In the U-Mart acceleration experiment toolkit, 4 price movement patterns are the referential spot price time series:

1. ascending, i.e., consecutively moving *upward* with some noise.
2. descending, i.e., consecutively moving *downward* with some noise.
3. reversal, i.e., firstly descending and then ascending like the semicircle with some noise
4. oscillating, i.e., periodically oscillating with some noise

Therefore, we will obtain the generated price movements according to the above 4 price movement patterns. As shown in Figs. 9.6, 9.7, 9.8 and 9.9, we see that these generated price time series will realize almost the referential spot price time series. However, we need to find some comparison measure to judge the similarity of the generated price time series to the referential price time series.

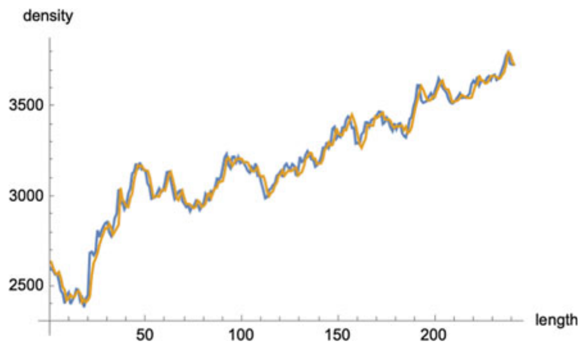


Fig. 9.6 Generated price time series in the minimally sized Nakajima-Mori agent composition given an ascending spot price time series

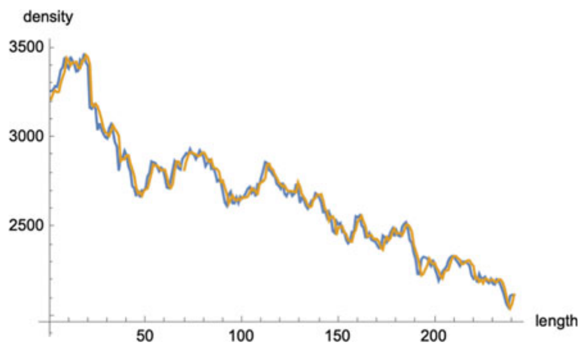


Fig. 9.7 Generated price time series in the minimally sized Nakajima-Mori agent composition given an descending spot price time series

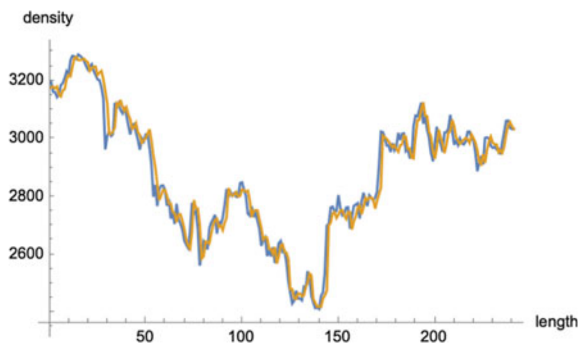


Fig. 9.8 Generated price time series in the minimally sized Nakajima-Mori agent composition given a reversal spot price time series

9.4 The Length Frequency Distribution Hinted by the Turing Machine

9.4.1 Fully Random Iterated Cellular Automata

In Aruka et al. (2021), it was suggested that a new index hinted at **the fully random iterated cellular automata**(FRICA) be derived, which has been argued in the

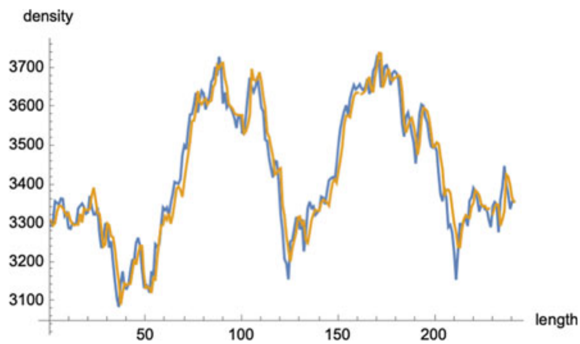


Fig. 9.9 Generated price time series in the minimally sized Nakajima-Mori agent composition given a descending spot price time series

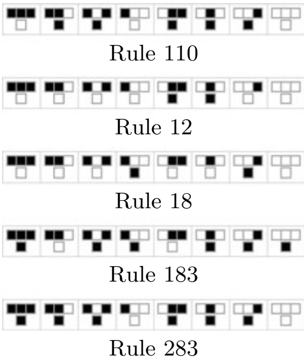


Fig. 9.10 Binary-colored presentation of rules 110, 12, 18, 183, and 238

heterogeneous rule dynamics of the **Turing machine**. Suppose that the rules of the CA are following 5 rules (Fig. 9.10):

110, 12, 18, 138, 238

We chose a rule randomly from this rule set, and this procedure is repeated consecutively. We are interested in observing how this CA will evolve in the plane. In Fig. 9.11, we apply the CA to run 50 rounds to plot the next shape[17]:

9.4.2 *Digit Length Distribution*

We then introduce the idea of **digit length distribution**. We measure the shapes generated by the FRICA vertically, i.e., from the origin (the initial round) up to the bottom (the final round) . We can find any sequence of the same digit (black or white color) alternately occurring in the above digit sequences. If we have a consecutive sequence of all 0, { 0,0,0 }, we say **length** of 0 is 3, and we also find the

[17] Black square: 1; white square: 0

Fig. 9.11 An evolution of the FRICA based on the rule set { 110, 183, 18, 238,12 } for **50 rounds**

Table 9.2 Stationary state transition

round	State of cells
t	0100
t+1	0100
t+2	0100
t+3	0100
t+4	0100
t+5	0100

Table 9.3 Length evolution of each cell column of the stationary state transition

	Length		
Column	Time development	0	1
1st column	00000	(5, 1)	NA
2nd column	11111	NA	(5, 1)
3rd column	00000	(5, 1)	NA
4th column	00000	(5, 1)	NA
length frequency each digit		(5, 3)	(5, 1)
Systemic length dist.		{(5, 4)}	NA

$*(x, y)$ means (the length, the number of repetitions), respectively

same subsequence from the whole sequence, which we describe (3, 2). The second component of 2 represents its **frequency.**

Simplest case It is well-known that the constancy, periodicity, and/or aperiodicity around the elementary cellular automata under consideration is identified with the patterns varying on one-dimensional space. According to our proposal, **the digit length** will be exposited in the following manner. First, we take the simplest numerical example of the stationary state transition. Take initially at t a sequence like 0100. Then, let the time development from t to $t + 5$ to be the series described by Table 9.2.

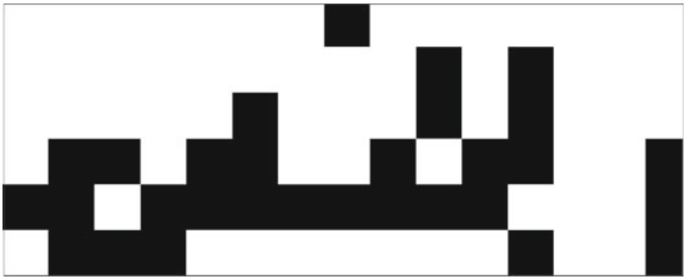


Fig. 9.12 Evolution of the FRICA based on the rule set {110, 183, 18, 238,12} for **5 rounds** ∗ In this case, we observe all the evolution of both left-hand and right-hand sides

Table 9.4 The state transition of FRICA

Round	The state of cells
t	000000010000000
t+1	000000000101000
t+2	000001000101000
t+3	011011001011001
t+4	110111111110001
t+5	011100000001001

Case of FRICA based on the rule set {110, 183, 18, 238,12} We then closely examine the earlier evolution of the FRICA based on the rule set {110, 183, 18, 238,12} shown in Fig. 9.12 (first 5 rounds) (Table 9.3).

Beginning with the same initial state, let then the time development from t to $t + 4$ to be the series described by Table 9.4.

We can easily calculate the length distribution of each digit, i.e., white and black as shown in Table 9.5. For simplicity, we only observe the evolution of the left-hand side from the 1st column to the 8th column, from where the initial cell starts.

Thus, we produce the frequency distribution of consecutive digits along each cell vertically. In our simulation, we first tentatively simulate 1,000 steps on 1,000 cells initially given. Figures 9.13 and 9.14 show a digit length frequency distribution **0/1**, for instance, at **the third column**.

9.5 Market Strategy Composition and Its Effect

Our purpose is to scientifically judge the similarity of the generated price time series to the referential price time series, as stated in Sect. 9.3.2. We are ready to measure a changed effect when changing the market strategy composition by the use of the digit length distribution, as argued in the previous section.

Table 9.5 Length evolution of each cell column of the FRICA

		Length	
Column	Time development	0	1
1st column	000010	{(4, 1),(1,1)}	{1,1}
2nd column	000111	(3, 1)	(3, 1)
3rd column	000101	{(3, 1),(1,1)}	(1, 2)
4th column	000011	(4, 1)	(2,1)
5th column	000110	{(3, 1),(1,1)}	(1,1)
6th column	001110	{(2, 1),(1,1)}	(3, 1)
7th column	000010	{(4, 1),(1,1)}	(1, 3)
8th column	000110	{(3, 1),(1,1)},	(1, 2)
Length frequency each digit		{(4, 2), (3, 2),(2, 1), (1, 6)}	{(3, 2), (2, 1), (1, 7)}

$*(x, y)$ means (the length, the number of repetitions), respectively

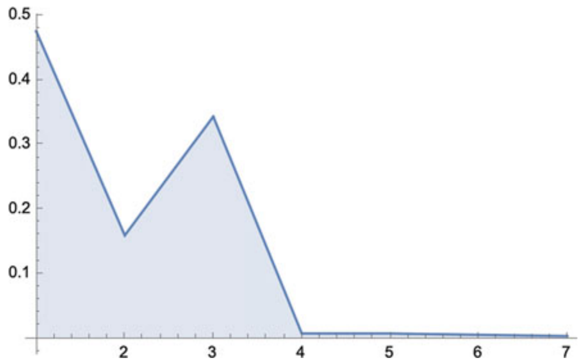


Fig. 9.13 Digit length frequency of **0**

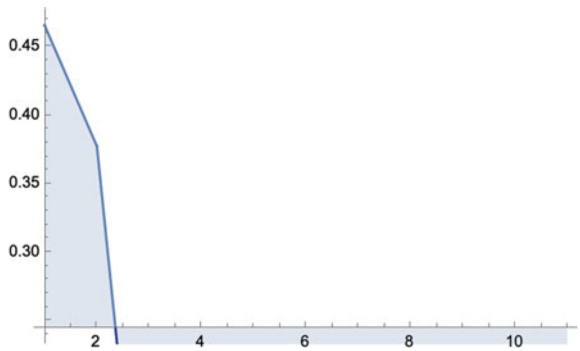


Fig. 9.14 Digit length frequency of **1**

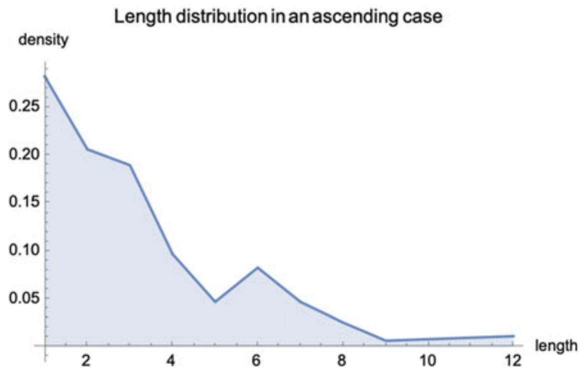


Fig. 9.15 Length distribution in the minimally sized Nakajima-Mori agent composition given an ascending spot price time series

Before proceeding this procedure to compare the *ex post* and the *ex ante* state, we make several comments. Our agents are irrelevant to a utility maximizing agent without minding the others' reactions of others. Our agent always follows a certain strategy chosen by himself/herself in advance. His/her decision will reflect at least partly the world reality. Thus, any divergence from changing the new agent set will then confirm a "rigged" outcome. In the above sections, first, we investigated the effects of manipulation in a market exchange. For convenience, we suppose an "agent strategy" with the aim of "manipulating" to attain the agent's agenda. In our case, the agenda is to realize the referential time series over time. It is natural for this strategy to operate it in a skillful manner, regardless of whether its operation succeeds or fails. Second, if the market should be contaminated by some manipulating force, then how a countervailing force be formed to cancel the contamination will be of great importance. Here, we are tempted to examine the effect of **neutralizing strategy**.

In Sect. 9.3.1, we examined by visual observation the similarity in the case of the minimum set added to 10 bodies of the **SFSpreadStrategy**. Indeed, as Figs. 9.6 and 9.9 show, the changed minimum configuration added by 10 bodies of the SFSpreadAgent cannot rig our agenda. Now, we apply our idea of the digit length frequency distribution to these cases. We first transform the price time series into the digital binary world of {0, 1}. That is, we allocate the variation to 1 when price behaves in an upward manner, otherwise, we allocate it to 0. Thus, this transformation will be of the same kind of volatility as this time series. Hence, the degree of similarity to its referential time series will be explicitly measured by using the binary transformed time series generated by a changed strategy composition. Practically, we will have two binary-transformed time series on the ex ante and the ex post states.

Figures 9.15, 9.16, 9.17, 9.18, 9.19, 9.20 and 9.21 represent the digit length frequency distribution in the minimal size Nakajima-Mori agent composition, while Figs. 9.16, 9.17, 9.18, 9.19, 9.20, 9.21 and 9.22 represent the digit length frequency distribution in in the minimal size with 20 bodies of SFSpreadStrategy added. It can be easily seen from the comparison between both figures that each frequency distribution on each price movement pattern between the ex ante and the ex post strategy composition cannot be mutually fully superior. Thus, we will conclude that

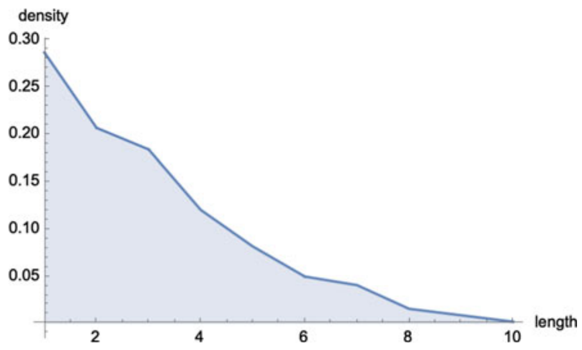


Fig. 9.16 Length distribution in the minimally sized added with 20 bodies of the SFSpreadStrategy given an ascending spot price time series

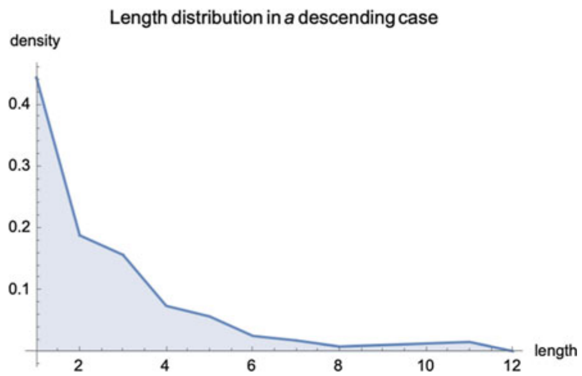


Fig. 9.17 Length distribution in the minimally sized Nakajima-Mori agent composition given a descending spot price time series

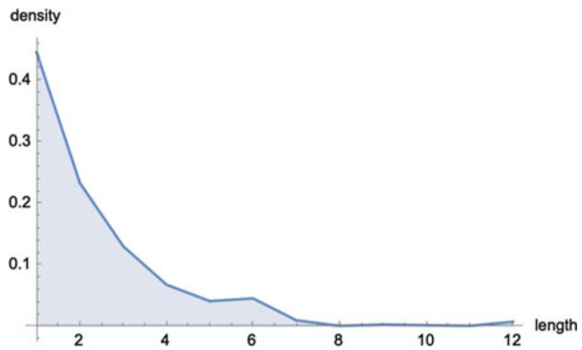


Fig. 9.18 Length distribution in the minimally sized Nakajima-Mori agent composition with 20 bodies of the SFSpread Strategy added given a descending spot price time series

the SFSpread strategy will have some *neutralizing* effect on the referential price time series.

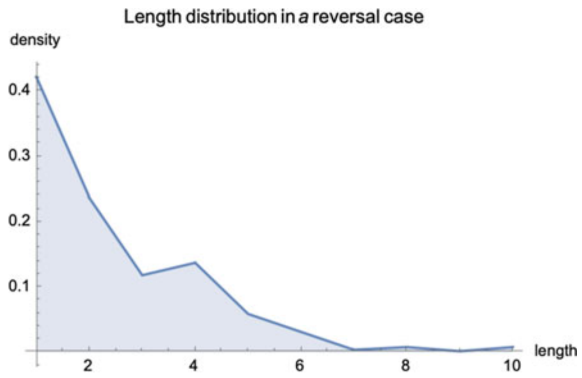


Fig. 9.19 Length distribution in the minimally sized Nakajima-Mori agent composition given a reversal spot price time series

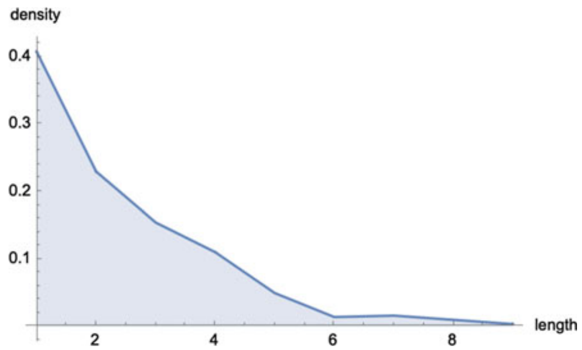


Fig. 9.20 Length distribution in the minimally sized Nakajima-Mori agent with 20 bodies of the SFSpread Strategy added given a reversal spot price time series

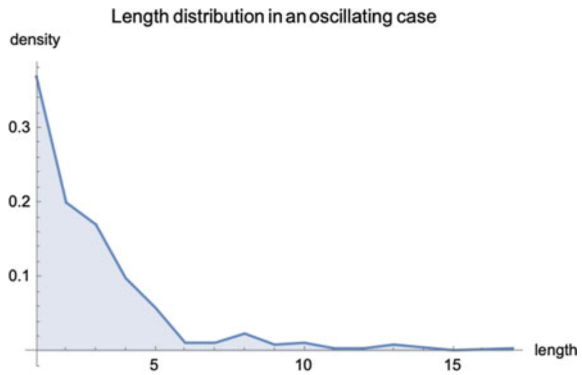


Fig. 9.21 Length distribution in the minimally sized Nakajima-Mori agent composition given an oscillating spot price time series

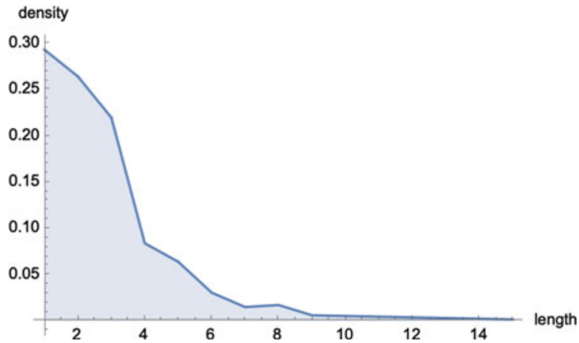


Fig. 9.22 Length distribution in the minimally sized Nakajima-Mori agent composition with 20 bodies of the SFSpread Strategy added given an oscillating spot price time series

References

Aruka Y (2015) Evolutionary foundations of economic science: how can scientists study evolving economic doctrines from the last centuries? (Springer series: Evolutionary economics and social complexity science, vol 1). Springer, Tokyo

Aruka Y, Nakajima Y, Mori N (2019) Evolut Instit Econ Rev 16(1):19–42

Aruka Y, Nakajima Y, Mori N (2020) The evolution of the exchange process: from the decentralized to the distributed digital exchange. Evolut Instit Econ Rev 17(2):379–398

Aruka Y, Nakajima Y, Mori, N (2019) The minimum heterogeneous agent configuration to realize the future price time series similar to any given spot price time series in the AI market experiment. In: Bucciarelli E, Chen SH, Corchado J (eds) Decision economics: complexity of decisions and decisions for complexity. DECON 2019. Advances in intelligent systems and computing, vol 1009, pp 85–92. Springer, Cham (2020)

Aruka Y, Nakajima Y, Mori N (2021) The minimum heterogeneous agent configuration to realize the future price time series in the AI market experiment: an updated version

Bernays E (1928) Propaganda. Horace Liveright, New York

Crandall JW, Oudah M, Tennom, Ishowo-Oloko F, Sherief Abdallah S, Bonnefon JF, Cebrian M, Shariff A, Goodrich MA, Iyad Rahwan I (2018) Cooperating with machines. Nat Commun 9(233)

Crowe J, Barbora J (2016) What's wrong with cartels? Federal Law Rev 401

Chartier G (2009) Economic justice and natural law. Nature of technology. Cambridge UP

Islam R (2021) Crypto-currency flow among bitcoin users: de-anonymization and dynamics of persistently-active big players by network analysis. Graduate School of Simulation Studies University of Hyogo, PhD thesis

Kita H, Taniguchi T, Nakajima Y(eds) (2016) Realistic simulation of financial markets analyzing market behaviors by the third mode of science (Springer series: Evolutionary economics and social complexity science, vol 4). Springer, Tokyo

Morishima M (1984) Cambridge UP, Cambridge

Nakamoto S (2019) Bitcoin: a peer-to-peer electronic cash system

Shiozawa Y, Nakajima Y, Matsui H, Koyama Y, Taniguchi K, Hashimoto F (2008) Artificial market experiments with the U-mart system (Springer series on agent based social systems, vol 4). Springer, Tokyo

Chapter 10
Market Power and the Hayek Hypothesis: An Experimental Investigation

Chung-Ching Tai and Bin-Tzong Chie

Abstract The inconsistency in the literature regarding price behavior in markets with market power suggested that traders are heterogeneous in their ability to discover and exercise their influences on market price. To examine whether inequality in subjects' market performance can be attributed to the heterogeneity of their cognitive capacity, we adopt a single-subject experimental design and inspect whether subjects can discover and exercise their market power by withholding part of their units. We use the Cognitive Reflection Test to gauge subjects' cognitive ability and investigate whether subjects with higher cognitive abilities will be better at exercising their market power to influence the price to gain extra profits. We observe heterogenous price behavior. However, there is no evidence that higher CRT scores were associated with better price convergence to the competitive equilibrium. Further, the revelation of information only altered some subjects' behavior.

Keywords Hayek hypothesis · Market power · Double auction · Cognitive ability

10.1 Introduction

How prices are formed in markets is the foundation of many economic theories, such as market equilibrium in macroeconomic models. Due to its profound implications, experimental economists have used experiments to validate competitive market theories for a very long time (Roth 1995). The Hayek Hypothesis, which proposes that market prices will converge to the competitive equilibrium, attracted much attention from experimental economists, and an abundance of studies has piled up since 1960s.s. A large part of this literature confirmed the Hayek Hypothesis in most conditions. However, we also observe a lot of inconsistent results, and sometimes

C.-C. Tai (✉)
Southampton Business School, University of Southampton, Southampton, UK
e-mail: c-c.tai@soton.ac.uk

B.-T. Chie
Department of Industrial Economics, Tamkang University, New Taipei City, Taiwan

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_10

deviations from the Hayek Hypothesis. While prior discussions of Hayek Hypothesis mainly focuses on *pricing mechanisms* and *market structure*, we are going to examine those deviations from an different point of view: we want to know whether people's *cognitive characteristics*, namely their cognitive ability and styles of cognitive processes, has any observable influence on how they behave in auction markets, and whether we can make use of this knowledge to explain market dynamics.

In what follows, we will briefly introduce experimental efforts dedicated to validate the Hayek Hypothesis first (Sect. 10.1.1), followed by evidence of deviations from competitive equilibrium occurred in the lab (Sect. 10.1.2). In Sect. 10.1.4, we will go through the literature on how cognitive ability influences economic performance. As we will see later in that section, the latest experimental results has clearly indicated that cognitive characteristics has great influence on people market behavior.

10.1.1 The Hayek Hypothesis and the Experimental Evidence

The 1974 Nobel laureate in economics, Friedrich Hayek, proposed that market can effectively aggregate unorganized and dispersed knowledge in the society and that the price system is in fact a mechanism for communicating information (Hayek 1945). The most striking fact about the price system is the economy of knowledge with which it operates: individuals need to know very little to make the right decisions. Prior to Hayek, Adam Smith also pointed out that the attainment of equilibrium allocations does not require any individual participant to have knowledge of the market as an allocation system (Smith 1937). Vernon Smith, being the 2002 Nobel laureate in economics, proposed the well-known Hayek Hypothesis (Smith 1982):

> Strict privacy together with the trading rules of a market institution are sufficient to produce competitive market outcomes at or near 100% efficiency.

The Hayek hypothesis emphasises that no explicit communication is needed to reach an efficient allocation of resources. Since then, experimental economists have tested various kinds of market institutions, including a variety of market structure and trading mechanisms (Holt 1995).

Being one of the common trading mechanisms for goods, double auction has been intensively tested since Smith (1962)'s pioneering work. Smith (1962) demonstrated that the actual transaction prices converged to competitive equilibria in a very fast pace. It turns out that convergence to the competitive equilibrium in double auctions is a robust phenomenon which has been observed over "at least a thousand sessions in a variety of designs" (Holt 1995). For example, Miller et al. (1977), Williams (1979), Smith (1982), Williams and Smith (1984), Davis et al. (1993), and Cason and Friedman (1993; 1996) have demonstrated that prices can even converge to competitive equilibria in dynamics or even random market conditions, although convergence in extremely nonstationary markets could be slower.[1]

[1] Compared to double auctions, convergence in other trading institutions were not as good as those in DA market experiments. See Smith (1964), Smith (1967), Williams (1973), Plott and Smith

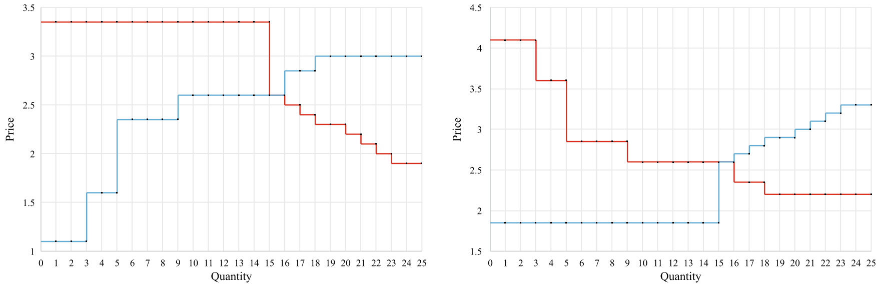


Fig. 10.1 Markets with the SMP design (the left part) or the BMP design (the right part). *Source* Constructed with data from Holt et al. (1986)

Despite so, there are actually instances where prices deviate from competitive equilibria, especially when some market participants have market power (Holt et al. 1986; Smith and Williams 1990; Davis and Williams 1991).

10.1.2 Traders with Market Power

In economics, market power means *the ability (of sellers) to affect prices*. For example, if buyers have market power, they will try drive down the prices to increase their benefits from the transactions. Holt et al. 1986 (hereinafter HLV86) conduct double auction experiments in which the demand-supply schedules grant some buyers or sellers the capacity to manipulate the prices, as shown in Fig. 10.1.

According to HLV86's design, some buyers or sellers can influence the market prices by withholding part of their units. For example, in the left part of Fig. 10.1, the market consists of five buyers and five sellers, and each of which has five units to buy (sell). The seller(s) can create a leftward shift of the supply curve and therefore drive up the price from 2.60 to 2.85 by withholding two units. We call this type of market the seller market power (hereinafter SMP) design. Similarly, in the right panel of Fig. 10.1, if the buyer(s) withhold two units, the market price will drop to 2.35 from the original equilibrium of 2.60. We call this type of market the buyer market power (hereinafter BMP) design. In HLV86's design, only two out of the five buyers (sellers) have the incentives and definite market power to influence prices. Manipulation could be achieved through a single buyer's (seller) unilateral action, or through coordinated behavior from two buyers' (sellers').

HLV86 conducted five SMP and two BMP experiments. Although they still observed convergence towards the original competitive equilibria, four of the exper-

(1978), Hoffman and Plott (1981), Smith and Williams (1982), Smith et al. (1982), Ketcham et al. ((Ketcham et al., 1984)), Davis and Williams (1986), Plott (1986), Walker and Williams (1988), Cason and Williams (1990), Davis and Williams (1991), Davis et al. (1993), and Kagel (2004) for examples.

imental markets ended up with constant deviations from the C.E. price. They further investigated into the differences among these experiments and found that the key is whether subjects with market power could discover and exercise the manipulation strategy. They noticed that experienced subjects tend to execute the market power more than inexperienced ones.

HLV86's findings were examined in a very similar design by Davis and Williams (1991). In Davis and Williams (1991)'s design, there are five buyers and five sellers in a market with five units to trade each. Both Seller 1 and Seller 2 had the market power if they withhold two units at the C.E. price. They observed deviations from the C.E. price in all their double auction markets.

10.1.3 How Should We View These Inconsistent Results?

Inevitably, experimental economists will observe individual differences among their subjects or even experimental sessions. How should we perceive such inconsistency? Generally speaking, we can ignore these differences and report the average behavior as long as these differences do not have any qualitative implications. For the aforementioned results, however, it is dangerous to discount those inconsistent results since they either supported or opposed the Hayek Hypothesis–you can't average them and claim that you have a mediocre support. Instead, as scientists we should look for boundary conditions where we can confidently validate the Hayek Hypothesis. The existence of such inconsistent results simply suggests that there are important factors left behind in the scientific process of pursuing the truth.

As HLV86 suggested, the information and institutional conditions of an oral double auction are not sufficient to ensure the convergence of prices to the competitive level in a standard-sized experiment. Although HLV86 suggested that experiences play this important role, Davis and Williams (1991) showed that experiences did not play the crucial role in explaining whether subjects exercised market power and caused the prices to deviate from C.E.

If it is not experiences, what else can explain the individual differences? We propose that cognitive characteristics (including cognitive ability and styles of cognitive processes) could be a plausible factor capable of explaining the aforementioned inconsistency. In the following session, we will go through a series of empirical and experimental evidence to support our conjecture. Most importantly, recent experimental studies have provided clear linkage between subject heterogeneity in cognitive characteristics and individual differences as well as discrepancy between experimental sessions.

10.1.4 Cognitive Ability and Economic Behavior

The relationship between cognitive ability and economic behavior has drawn increasing attention from experimental economists. The interests of using controlled experiments to study the effects of cognitive capacity on economic behavior starts with the classical problem of Prisoner's Dilemma. In their experimental repeated prisoners' dilemma games, both Segal and Hershberger (1999) and Jones (2008) found that higher cognitive ability were associated with more cooperative behavior. Jones (2008), specifically, found that subjects cooperate more often by 5–8% whenever their SAT (Scholastic Aptitude Test) scores are 100 points higher.

In a context of a sequential Prisoner's Dilemma game, Burks et al. (2009) discovered that subjects with higher IQ can predict their opponents' behavior more accurately, be they the first movers or the second movers. Likewise, Devetag and Warglien (2003) employed various kinds of game context and found that the larger the subjects' short-term memory, the closer their decisions are to the rational choices. This is an interesting results, however, the most inspiring finding comes from their later work. Devetag and Warglien (2008) use a delicate design to test whether subject "misunderstand" the game before they proceed to make their decisions. The results intriguingly subvert one's expectations–for some complex games, subjects with less short-term memory tend to misrepresent the games more and act rationally upon these "simplified" games.

Cognitive Ability and Market Performance

It is natural to ask, after reviewing similar questions in various domains mentioned above, whether *inequality in agent's market performance can be attributed to the heterogeneity of their cognitive capacity*. By comparing subjects' actual bids to the theoretical break-even bids, Casari et al. (2007) examined whether subjects suffer the winner's curse in common value auctions. They found that those whose SAT/ACT (American College Test) scores are below the median are more susceptible to the winner's curse.

A similar but slightly different measure of people's cognitive ability is the Cognitive Reflection Test (CRT). The CRT was proposed by Frederick (2005), and it is a three-item test to characterize people's styles of cognitive processes. To be precise, the test can be used to distinguish people who tend to use System 1 thinking from those who use System 2. System 1 is a spontaneous thinking process with little conscious deliberation, while System 2 thinking consumes more attention and is more reflective (Kahneman and Frederick 2002; Stanovich and West 2000).[2] What does CRT has anything to do with economic behavior? A series of studies have shown that

[2] Although CRT is designed to distinguish styles of cognitive processes, Frederick (2005) demonstrated that CRT scores are correlated with people's time preference, risk preference, and other cognitive measure such as SAT/ACT/Wonderlic Personnel Test. It turns out that CRT has become a very popular tool for experimental economists as a proxy of cognitive ability.

CRT scores are negatively correlated with many decision bias, such as the conjunction fallacy, illusion of control, overconfidence, base rate fallacy, and conservatism (anchoring effect) (Oechssler et al. 2009; Bergman et al. 2010; Hoppe and Kusterer 2011; Noori 2016). As for people's behavior in game-theoretic environment, experimental evidence shows that subjects with higher CRT scores are more likely to play dominant strategies in the beauty contest game (Schnusenberg and Gallo 2011; Brañas-Garza et al. 2012) and bank-run games (Kiss et al. 2016), accept unfair offers in the ultimatum game (Calvillo and Burgeno 2015), and trust more in trust games (Corgnet et al. 2016).

For this chapter, the most important thing which can be learned from the literature is how CRT performance relates to subjects' behavior in market experiments. Corgnet et al. (2015) followed Smith et al. (1988)'s asset market design but with certain dividends (Porter and Smith 1995), and motivated subjects in two different schemes–earned money vs. house money. They found that no matter how the endowments were given, the higher the subjects' CRT scores, the more money they earned. To be more specific, subjects with lower CRT scores tended to buy shares when prices were above fundamental values, and sell share vice versa.

Breaban and Noussair (2015) had two treatments: bearish market and bullish market. For each treatment, subjects played two consecutive 15-period market experiments (Market 1 and Market 2), and that makes subjects experienced in Market 2. According to subjects' trading behavior, they were classified into three categories: *Fundamental Value* traders, *Momentum* traders, and *Rational Speculators*. Breaban and Noussair (2015) found that there is a significant correlation between CRT scores and being a fundamental value trader in Market 1, and the greater the average CRT score, the smaller the differences between market prices and fundamental values.

The most interesting research regarding CRT and market performance comes from Noussair et al. (2016)'s futures market experiments. In their study, there are both a spot market and a futures market for participants to trade. Their research goal is to know whether the existence of a futures market can help price convergence in the spot market. They had the same experimental designs in two locations, and observed different price patterns.[3] Noussair et al. (2016) found that the average CRT score of a trader cohort is negatively correlated with the degree of mispricing when no futures market is present, and this relationship disappeared if a futures market exists. However, from an individual point of view, traders with higher CRT scores had greater earnings no matter a futures market was presented or not. In additional to individual differences, they found inconsistent results from the data collected in the two locations–if measured in relative absolute deviation (RAD) and volatility, the bubble in one location is significantly greater than that in another. It turns out that the subject pools in these two locations had different CRT scores, and differences in their CRT scores can entirely explain the difference in the RAD and volatility measures between the two locations where the experiment was conducted.

[3] Their experiments took place in Waikato University, New Zealand and Tilburg University, the Netherlands.

10.1.5 Research Questions and Hypotheses

Noussair et al. (2016)'s study has two important features. First, they found that cognitive characteristics can explain differences in behavior–in the level of buying/selling actions as well as their final earnings. Second, they could explain discrepancy between experimental sessions, and that discrepancy is very hard to explain by other factors because the experiments had the same designs, parameters, and even had roughly the same amount of subjects. Their findings, as well as Corgnet et al. (2015)'s and Breaban and Noussair (2015)'s discoveries, gave us a strong hint that cognitive characteristics is something we cannot ignore when we observe inconsistent results for the Hayek Hypothesis.

Although a large literature of agent-based simulation studies have focused on how agents' level of rationality impact the Hayek Hypothesis, so far we haven't seen any systematic study using human experiments to answer the same question.[4] In this chapter, we aim at providing fundamental and direct evidence of how people's cognitive characteristics cause them to act differently and therefore impact price dynamics.

More specifically, the first goal of this study is to observe whether cognitive ability can explain the inconsistency among individual traders:

Hypothesis 1 *Subjects with higher cognitive abilities will be better at exercising their market power to influence the price to gain extra profits.*

Friedman and Ostroy (1995) proposed that under certain parametric structures, market participants may have the ability to influence market price favorably. As mentioned above, HLV86's experiments also showed that some of the subjects did succeed in exercise market power and gained extra profits. To exercise market power, subjects has to grasp the demand-supply structure or at least the pattern of marginal units first to realize that they have market power. Since the demand-supply structure is not public information, subjects can only learn the structure by observing past market information. Our conjecture is that subjects with higher cognitive abilities will perform better in eliciting this information from experiences or historical data, and therefore have greater chances to exercise market power.

Hypothesis 2 *Markets consisting of participants with higher cognitive abilities will converge more quickly to the competitive equilibrium than those consisting of participants with lower cognitive abilities.*

From Corgnet et al. (2015), Breaban and Noussair (2015), and Noussair et al. (2016)'s experimental studies, it is evident that subjects with higher cognitive abilities will behave more like fundamentalists. An indirect but parallel conjecture is that in auction markets, the market price will converge more quickly to C.E. if market participants have higher cognitive abilities.

[4] See Gode and Sunder (1993) and Duffy (2006) for some discussions on how agent-based computational economics uses zero-intelligence agents to study price convergence towards the competitive equilibrium.

10.2 Experimental Design

To answer our research question and examine our hypotheses, we need to have different experimental designs in our two-year project. A specific set of psychometric tests or simple tasks to detect subjects' personal traits and behavioral tendencies will also be performed. In this section, we will introduce the designs of our market experiments. Experimental procedures, illustrations of the tasks, and analytical strategies will also be provided in this section.

10.2.1 The Double Auction Experiments

The aims of this experiment is to understand how subjects with different degree of cognitive abilities or styles of mental processes conduct their bidding behavior in auction markets.

Following Easley and Ledyard (1993)'s terminology, in our auction market, the *true values* of the commodities are given to buyers as integers and are ranked $V^1 \geq V^2 \geq \cdots \geq V^n \geq 0$, where V^i is the ith highest value and there are n units for all the buyers to purchase. A buyer b will be assigned a subset of these units $V_{b1} \geq V_{b2} \geq \cdots \geq V_{bB}$ and will trade them one at a time in the sequence $b1, b2, \ldots, bB$. No recontracting is allowed.

Similarly, the *true costs* given to sellers are also integers and are ranked as $0 \leq M^1 \leq M^2 \leq \cdots \leq M^m$, where M^j is the cost of the jth unit and there are m units for all the sellers to sell. A seller will be assigned a subset of these units $M_{s1} \leq M_{s2} \leq \cdots \leq M_{sS}$ and will trade them one at a time in the sequence $s1, s2, \ldots, sS$. No recontracting is allowed.[5]

[5] The bidding behavior of a buyer agent can be summarized as follows:

1. The agent collects market information and defines the following prices.
    - Among all current buyer bids, the highest bid and transaction price, whichever is higher, is defined as the *highest price*.
    - Among all current seller bids, the lowest offer and transaction price, whichever is lower, is defined as the *lowest price*.
    - The lowest bid of all buyers in the previous period is defined as the *buyer's convergence price*.
    - The highest offer of all sellers in the previous period is defined as the *seller's convergence price*.
    - The buyer agent's own *current true value* is the highest one of the holding true values.
2. The buyer agent's *bidding price* is ranged between its own *current true value* and the above-mentioned established price.
3. The buyer agent's *bidding price* shall be the lowest of the *current true value*, the *highest price*, and the *buyer's convergence price*.
4. If the *bidding price* is the *highest price* and there is no transaction, the *highest price is increased by one*, and then substituted into the former step to determine the *updated bidding price* in the next second.

Each buyer (seller) knows only her own values (costs) and this is no other public information about the demand-supply schedules. The number of buyers and sellers as well as past bids, offers, and prices are public information. We plan to have four buyers and four sellers in the double auction market.

Participants in a double auction market are in fact facing a very complex decision problem in the sense that they have to decide when to bid, how much to bid, and whether or not to accept current offers or not (Easley and Ledyard 1993). Because the subjects cannot know the effect of their actions on others' actions, there exists a complex and interactive decision problem with incomplete information.

In our double auction experiments, in order to clearly observe individual subject's bidding behavior and its effect on market price, we have to control the environment and bypass the complicated entanglement of strategic interactions. To achieve this goal, we have several special design features in our double auction experiments as follows:

1. Each human subject play against a set of computer robots instead of other human subjects.
2. Human subjects all play the same role in the market.

Feature 1 is to avoid complex strategic interactions among human participants. When subjects are facing the same set of computer agents, they are facing the same set of strategies. In this case, we as experimenters not only can know exactly the strategies employed by other robots so as to analyze human subjects' behavior, we can also manipulate robots' strategies to discuss how human subjects respond in different strategic environments. In addition to the above advantages, when subjects are informed that they are playing against computers, experimenters can make sure that subjects will be motivated only by financial reward and not by other drivers such as altruism, reciprocity, fairness, etc., and the possible forms of sophisticated behavior induced by the belief that one's opponent may be irrational, revengeful, etc. is ruled out (Devetag and Warglien 2003).

The robot used in this study is the Easley-Ledyard bidding agents, which follows the bidding strategy proposed by Easley and Ledyard (1993). With this design, our human subjects will be able to make their decisions in an environment free from behavioral uncertainty resulting from other-regarding preferences.

Feature 2 is to make sure that every human subject has the same starting point. That is, every human subject will be assigned the same set of true values. We will also assign our subjects as buyers or sellers in a random way.

5. If the *bidding price* of the former step has not been executed, then the agent will take the lowest value among the *current true value*, the *highest price*, the *buyer's convergence price*, and the *transaction price increase by one* to determine the *updated bidding price* in the next second.

The seller will act in a similar way but from the opposite direction.

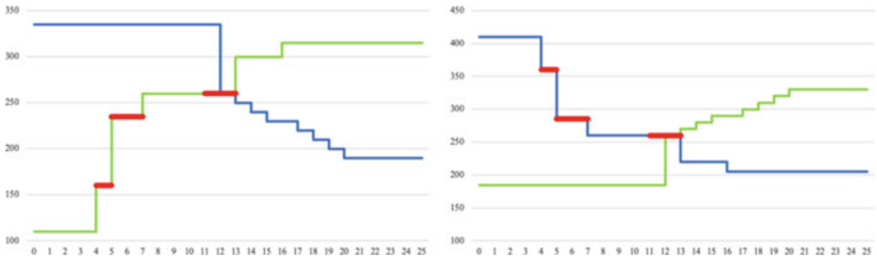


Fig. 10.2 Markets with the SMP design (the left part) or the BMP design (the right part)

10.2.2 Market Structures

Recall that Hypothesis 1 presumes that subjects with higher abilities will be better at exercising market power. To test whether Hypothesis 1 holds, we need a market structure which endows our subjects market power. We have both the BMP design and SMP design similar to HLV86's to facilitate our experiment. Figure 10.2 illustrates the demand and supply curves in our *asymmetric* double auction experiments.

In this market structure, there are several marginal units of which the true values are the same as the equilibrium price (the flat section of the demand curve right at the equilibrium price 260). We will endow our buyer and seller subjects five intra-marginal units, and two of them will has true values of 260. These five units are highlighted in red in Fig. 10.2. Take the BMP design for example, if this buyer subject withhold these two marginal units, the demand curve will be shifted leftward for two units and theoretically will result in a new equilibrium price of 220. Under this new equilibrium price, our buyer subject can earn extra profits than before.

Compared with HLV86's experiments, we believe that our experimental design will make it easier for our human subjects to discover their market power. The first reason is that we have only one human subject, and other market participants are robots. In this case, human subjects will be situated in a less complicated environment for several repetitive periods. A calculating human buyer is expected to learn that two of her intra-marginal units are in fact unprofitable, and abandon those units will yield extra bonus. In HLV86's experiments, other human subjects' and seller agents' actions were probably making things more complicated and not easy for the key buyers to discover their market power. Following this logic, we expect that subjects who gain more CRT scores (which means they are more likely to respond based on calculation rather than on gut feelings) will discover and exercise their market power more easily.

Treatments and Interface

To fully understand whether subjects have the ability to recognise their market power, our market experiments consist of two treatments. Treatment 1 is called the "Non-Informed" or *NI* Treatment. In this treatment, subjects only knew their own values and had no information about the market structure. There will be eight trading periods in this treatment, and each period will last for 150 s.

Treatment 2 is called the "Informed" or *I* Treatment. In this treatment, subjects will be given a sheet demonstrating the information about the demand-supply schedules. To be more explicit, the information in that demonstrating sheet include (1) detailed tables of unit values for all buyers and all sellers and (2) a demand-supply diagram as can be seen in Fig. 10.2. With such a demonstration, subjects should have all the necessary information to recognise their market power and therefore deduce the withholding strategy. Again, there will be eight periods in this treatment, and each period will last for 150 s.

In our experiments, every subject will first play the *NI* treatment for eight periods. After that, we will distribute the demonstrating sheet and explain the information in it. After the introduction of the market structure, the subjects will play the *I* treatment in the same market as the same role.[6]

The experiment was implemented using Z-Tree (Fischbacher 2007), and the trading interface is demonstrated in Fig. 10.3. There are several areas in this interface:

- The private information panel (the top-left panel): including the current profit and accumulated profit on the top and a table with (1) the value, (2) whether an order has been submitted, (3) whether it has been traded, and (4) the transaction price for each of the five units.
- The order submission panel (the middle-left panel): subjects can enter their limit price and press the red 'Submit' button. They can also press the grey 'Withdraw' button to cancel an outstanding order.
- The order book panel (the bottom-left panel): this table will reveal the best five buying prices and best five selling prices.
- The transaction history (the top-right panel): this table will show all the transaction information, including the buyer numbers, the seller numbers, and the transaction prices.
- The plot (the bottom-right panel): this is a graphical representation of the time series of the transaction prices.

10.2.3 Measuring Cognitive Ability

In this study, we use the Cognitive Reflection Test (CRT) to measure subjects' cognitive characteristics. The Cognitive Reflection Test measure how far people's mental processes are away from System 1 thinking. The higher the scores, the closer one's mental process to System 2 reasoning process. In this project, we adopt the traditional three-question CRT as follows, with the questions translated into Chinese for our subjects.

[6] However, this is a new experiments, and the Easley-Ledyard agents will restart their adopting behavior from scratch again.

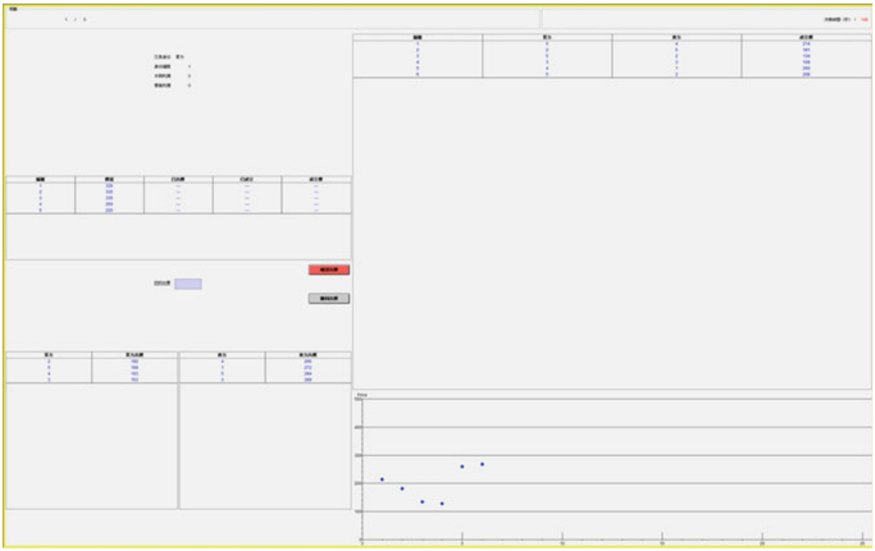


Fig. 10.3 The trading platform

1. A bat and a ball cost a total of 1.10 Euro. The bat costs 1 Euro more than the ball. How much does the ball cost?
2. If it takes five machines five minutes to make five widgets, how long does it take 100 machines to make 100 widgets?
3. In the lake there is a patch of lily pads, which doubles in size every day. It takes 48 days for the patch to cover the entire lake. How many days does it take the patch to cover half of the lake?

Translated version: The CRT will be carried out in Z-Tree (Fischbacher 2007).

Payments and Experimental Procedures

All the experiments were conducted at Tunghai University, Taiwan from June to August in 2019. The recruitment will be done via emails and website announcement. Subjects were motivated by real money, including a fixed amount of attending fee (NT$100) and a bonus calculated based on their performance in the double auction experiments. The exchange rate between ECU and NT dollars was 0.05. To encourage trading, subjects would be given a trading bonus for each transaction made. More specifically, they would receive NT$0.25 whenever a unit was bought or sold. The average bonus was NT$308.33, which means the bonus was three times the amount of the attending fee.

A total of 39 effective subjects were recruited. Among them, 20 have participated in the SMP experiments, while 19 participated in the BMP experiments. After subjects arrived in the computer lab, an introduction to the auction rules were given along with some trading examples. The subjects would play the *NI* treatments first, followed by the *I* treatment. After the auction experiments, subject then took the CRT tests as a measure of their cognitive ability.

10.3 Results and Discussions

In this section, we will respond to our research hypotheses about aggregate prices and individual performance. We will first examine whether market prices converge to the competitive equilibrium based on graphical evidence as well as quantitative measurements. Then, we will discuss the potential influences of subjects' cognitive ability by examining the relationship between their CRT scores and their market performance.

10.3.1 Market Power and Price Convergence

The most straightforward way to check whether subjects made use of their marker power is to examine the behavior of the market prices. Figure 10.4 demonstrates some typical price behavior observed in our *NI* experiments. In these experiments, subjects were only aware of their own private values, and they could only explore the market using trial and error. Our first question is therefore whether subjects were able to discover their market power and then manipulated the market prices. Here we call the original competitive equilibrium *CE* and the equilibrium when our subjects withdrew their units *MP*.

In Fig. 10.4, the top panels of both the SMP and BMP experiments are typical examples of fluctuating prices between CE and MP; the central panels are examples of convergence to CE; the bottom panels are examples of convergence to MP. From Fig. 10.4, it is clear that we have observed mixed results: not every subject can discover the market power.

To have a better classification about price behavior, we use net absolute deviation (NAD) as a measurement of deviation of a transaction price from a specific target price. To determine whether convergence has occurred, we make use of two types of NADs – NAD-CE and NAD-MP:

$$NAD - CE = Abs(P - CE); \tag{10.1}$$

$$NAD - MP = Abs(P - MP), \tag{10.2}$$

where P is the transaction price. We then compute the average NAD in every period. To compare the NADs across experiments and treatments, we further divide the original NADs by the CE price, which is 260 in all the markets. A price series will be regarded converging towards the CE if the average standardised NAD-CE in the last four periods (Period 5 to Period 8) is less then 5%. Similarly, a price series will be regarded converging toward the MP if the average standardised NAD-MP in the last four periods is less then 5%.

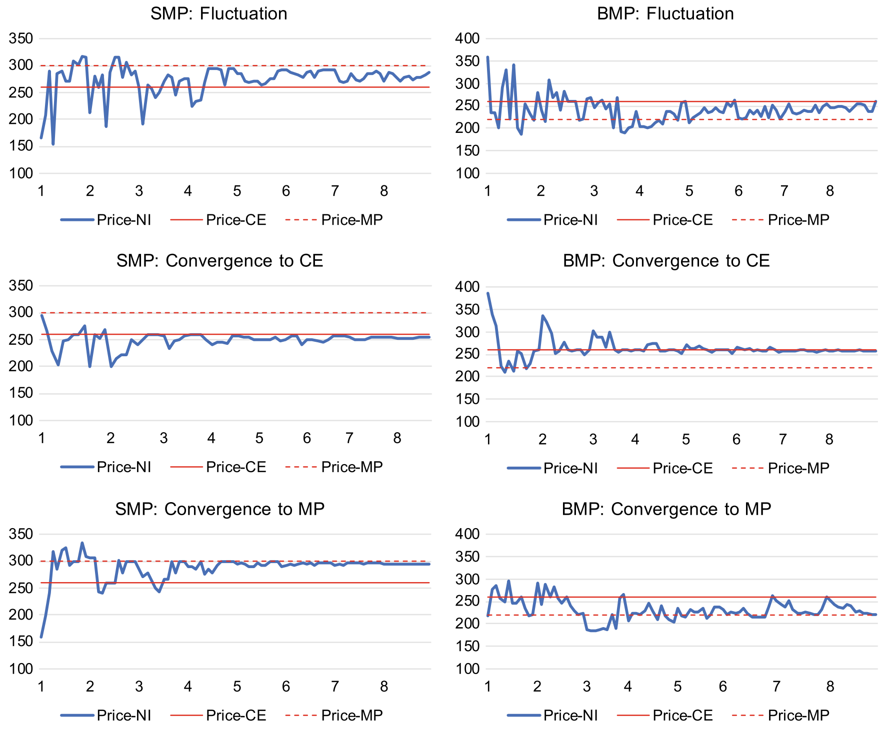


Fig. 10.4 Typical price patterns in SMP and BMP markets

Table 10.1 Classification of price patterns

Experiment	Pattern	Number of instances
NI-SMP	Convergence to CE	5
	Fluctuation	12
	Convergence to MP	3
NI-BMP	Convergence to CE	3
	Fluctuation	15
	Convergence to MP	1

According to this definition, there are three (out of 20) instances of market power in the *NI*-SMP experiments and only 1 (out of 19) instance in the *NI*-BMP experiments. Table 10.1 summarises the number of instances of the three types of price patterns–fluctuation, convergence to CE and convergence to MP. Since only a few subjects succeeded in exercising their market power, we then check whether cognitive ability has anything to do with the ability to recognise the market power.

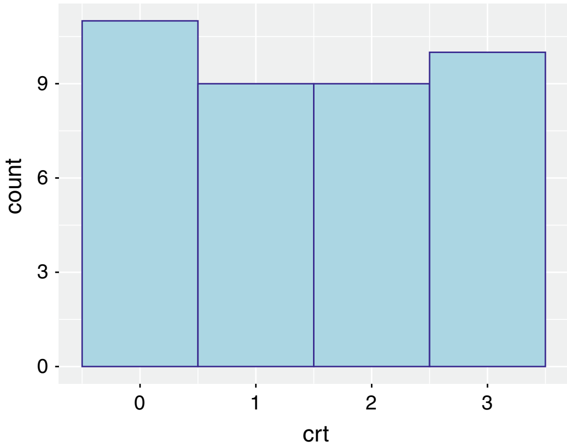


Fig. 10.5 Distribution of subjects' CRT scores

10.3.2 Cognitive Ability and Market Power

To see how subjects' cognitive ability influences their ability to exert their marker power, we first check the distribution of their CRT scores, as demonstrated in Fig. 10.5. It is clear from Fig. 10.5 that the distribution of CRT scores almost follows a uniform distribution for our 39 subjects, which means our subjects were quite different in their thinking styles. If that is the case, is it possible that those who succeeded in exercising their market power tend to think things differently from others?

To answer this question, we can use two methods. Method 1 is to check whether those who exercised market power had different CRT scores than others. Based on the classification of Table 10.1, we classify our subjects into three groups–those who lead to convergence to CE as well as MP and those who lead to fluctuation in market prices. Figure 10.6 demonstrates the CRT scores of these three groups. Surprisingly, those who exercised the market power appeared to be more impulsive–who rely more on System 1 thinking On the contrary, subjects with higher CRT scores seem to lead to fluctuating prices. The mean CRT scores for these three groups of subjects are 1.38 (CE), 1.59 (Fluctuation) and 0.75 (MP), respectively. Despite so, Kruskal-Wallis test shows no significant difference among these three groups (with p-value 0.387),

Method 2 is to check whether subjects have tried to withdraw their units so as to manipulate the market prices, as we know that subjects need to hold at least two units so as to move the demand or supply curves. The average units traded by our human traders in the last four periods in the *NI* treatments are 3.88 (CE), 4.03 (Fluctuation) and 3.81 (MP), respectively. Although subjects in markets of which the prices converged to MP did have fewer transactions, it appears that there are no significant differences among these three groups, either (with a p-value of 0.5163 in Kruskal-Wallis test).

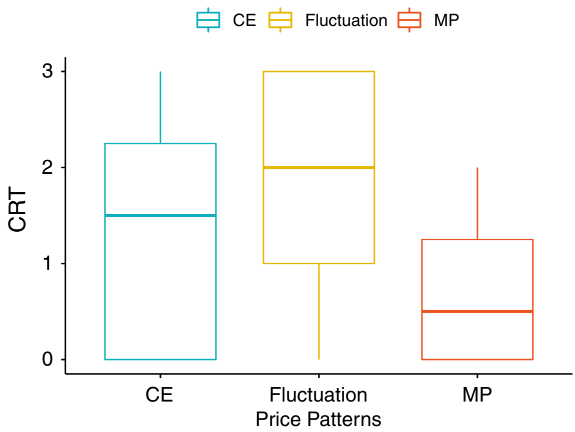


Fig. 10.6 CRT scores for subjects with different price patterns

Based on the above results, it appears that those who exercised market power were not withholding their units in the last four periods. Then what happened to those markets where prices converged toward the MP? Since we have only a few markets converging towards MP, we can scrutinise the detailed records to examine what has happened in those markets. Table 10.2 provides the detailed information about average market price, total number of trades and number of trades by the subjects for every period in the markets where prices eventually converge to MP.

We can see very clearly from this table that, whenever the subjects reduced their units at the early stage, the market prices would approach MP (300 for SMP and 220 for BMP) immediately. Once market price was driven up by subjects' withholding behavior, the prices will remain close to MP even if subjects increased their number of trades afterwards. To see whether this is what has actually occurred, Fig. 10.7 plots the numbers of trades from human subjects in the first four periods for each of the three price pattern categories. We can see very clearly that the more trades subjects tried to achieve, the closer the price to CE. On the other hand, subjects made fewer trades in those markets where prices converged to MP. Despite that this observation makes good sense, the statistical test shows that the differences among these three groups are not significant enough ($\chi^2 = 2.6621$, df $= 2$, p-value $= 0.2642$).

In sum, in this section we have shown that some of the subjects did withhold their units to manipulate market prices at early stages. Also, there is a no evidence that subjects with lower CRT scores tend to be the ones who exercised their market power. This observation is counter-intuitive as we thought that it would require more strategic thinking when a trader tries to exploit market opportunities. Despite so, this observation lacks statistical support probably due to the small sample size, especially when we only have 4 instances of convergence towards MP. We therefore need to reject our hypothesis *H1*, which assumes that Subjects with higher cognitive abilities will be better at exercising their market power to influence the price to gain extra profits.

Table 10.2 Average price and number of trades in markets converging to MP

Subject	SMP-4			SMP-8			SMP-19		
	Average price	Subject trades	Total trades	Average price	Subject trades	Total trades	Average price	Subject trades	Total trades
Period 1	281.75	1	22	270.27	5	28	260	3	23
Period 2	279.67	2	19	282.57	3	28	271	4	23
Period 3	275	2	22	298.38	3	26	297.45	1	21
Period 4	291.08	4	24	286.42	5	24	282.69	4	26
Period 5	293.83	4	24	287.62	3	23	285.23	3	24
Period 6	295.17	4	24	286.25	4	24	286.46	5	26
Period 7	295.25	4	24	290.92	4	24	284.38	5	26
Period 8	294.92	4	24	286.75	4	24	288.67	4	24
Subject	BMP-18								
	Average price	Subject trades	Total trades						
Period 1	251.08	3	23						
Period 2	253.58	4	24						
Period 3	205.2	2	18						
Period 4	222	2	24						
Period 5	227.09	3	22						
Period 6	226.92	3	26						
Period 7	234.46	3	24						
Period 8	233.08	4	24						

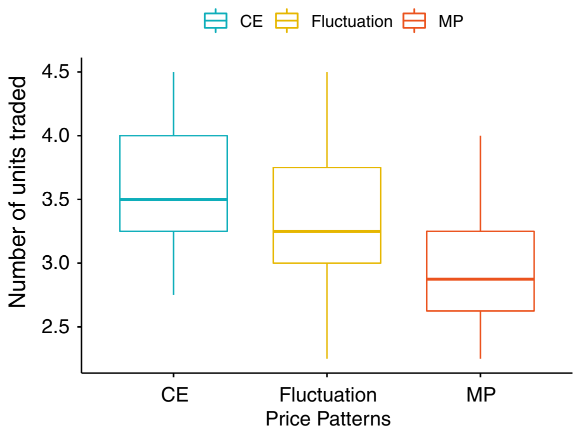


Fig. 10.7 Number of trades made by human subjects for different price pattern categories

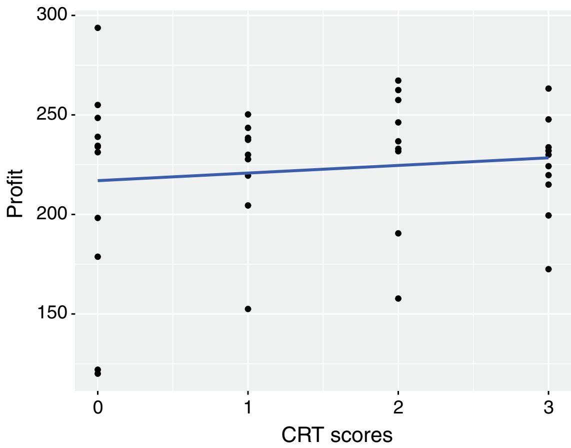


Fig. 10.8 The relationship between subjects' profit in the last four periods in the *NI* treatments and their CRT scores

10.3.3 Cognitive Ability and Market Performance

One of our research questions is whether subjects with higher cognitive ability will be able to gain more profits through their bidding behavior. As can be seen in Tai et al. (2018) that subjects with higher working memory capacity earned more profits in double auction markets, we intend to examine whether subjects with higher CRT scores would be more successful in our continuous double auction experiments as well. Figure 10.8 presents the relationship between subjects' profit in the last four periods in the *NI* treatments and their CRT scores. It appears that there is a weakly positive relationship between subjects' cognitive ability and their market performance.

We adopt the same analytical strategy as Tai et al. (2018) by dividing our subjects into two groups–the High group for those whose CRT scores is higher or equal to 2 and the Low group for those whose CRT scores are lower or equal to 1. We then examine whether the High group earned more profits than the Low group in the last four periods of the *NI* treatment. To distinguish the improvement in performance caused by learning and cognitive ability, we examine both the High and Low groups' performance in the first four period, where the learning effect may be weaker, and the last four period.

The results of the t test and the Wilcoxon Rank Sum test show that the two groups' average profits are not significantly different in both the first and the second halves of the *NI* treatment ($t = 1.9321$, df $= 32.486$, p-value $= 0.06211$ and $W = 247.5$, p-value $= 0.1092$ for the first four periods; $t = 0.78615$, df $= 33.567$, p-value $= 0.4373$ and $W = 203.5$, p-value $= 0.7149$ for the last four periods). Therefore, we couldn't find significant evidence supporting the view that subjects' market performance will be influenced by their thinking styles as captured by the CRT scores.[7]

[7] We also compare both groups' performance for the eight periods, and the difference is still insignificant.

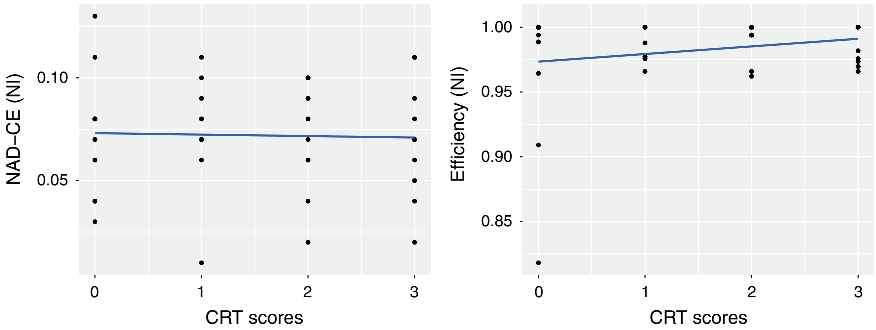


Fig. 10.9 The relationship between subjects' CRT scores with NAD-CE in the last four periods (left) and the market efficiency in the last four periods (right) in the *NI* treatments

10.3.4 Cognitive Ability and Convergence Towards the Competitive Equilibrium

According to experimental studies in the literature, asset prices tend to converge to the fundamental values when market participants had higher cognitive ability. This is imaginable since subjects had a chance to calculate the fundamental values based on their knowledge and the available information. In the *NI* treatment of our auction experiments, subjects only knew the values of their own units. In this circumstance, whether higher cognitive ability will help price convergence is an open question.

From Fig. 10.6, it appears that subjects who were in the CE category have lower CRT scores than those in the Fluctuation category, despite that the difference is insignificant (Wilcoxon rank sum test: $W = 96.5, p$-value $= 0.6552$). To better judge whether CRT scores are related to market convergence to CE, we employed two measures of convergence: NAD-CE and Market Efficiency. Market Efficiency is defined as the ratio of realised profits divided by the potential profit, which is the profits that should be realised when market participants are bidding in a truthful manner.

Figure 10.9 presents the relationships between subjects' CRT scores and the corresponding NAD-CE and Market Efficiency in the last four periods in the *NI* treatment. It is evident that there are no clear linear relationships existing in these plots. Simple regression analyses also suggest that CRT cannot serve as a significant explanatory variable for neither NAD-CE nor Market Efficiency (the adjusted R^2 for these two models are –0.02611 and 0.01854, respectively). Based on these results, we have to reject our hypothesis *H2*. This implies that subjects' CRT scores have no influences on the market prices in terms of the convergence towards competitive equilibrium.

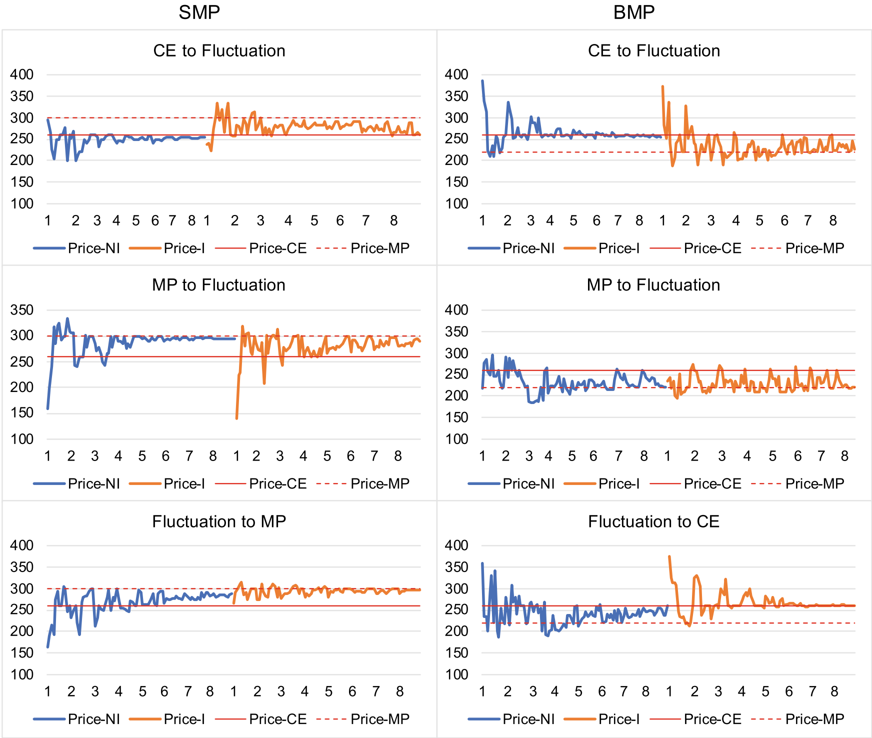


Fig. 10.10 Typical price patterns for *NI* and *I* treatments

10.3.5 Revelation of Market Information

One distinct feature of this study is that we allow subjects to have full information about the market in the *I* treatment. In this treatment, subjects were informed about the detailed values and shapes of the demand and supply schedules. Since they already had experiences in these markets, such an extra information should be very helpful in improving their trading strategies. Note that in our design, the Easley-Ledyard agents will be reset in the *I* treatment. If there are changes in price patterns, that should be due to changes in subjects' behavior.

With the introduction of additional information, how would subjects adjust their behavior? It turns out that subjects reacted quite differently when they had full information about the market environment. In terms of price patterns, some of the subjects adjusted their bidding behavior so the prices deviation from CE or MP in the *NI* treatment; others made use of this information and eventually reached the equilibrium at the CE or MP level. Figure 10.10 presents typical price patterns observed in our experiments.

Table 10.3 Profits and market efficiency after information revelation

	Profit			Market Efficiency		
	NI	*I*	*p*-value	*NI*	*I*	*p*-value
Mean	222.56	227.73	0.3846	0.98	0.99	0.2776

In the top-left plot, we can see that the price converged to CE in the first eight period (the *NI* treatment). After the revelation of the information, the human seller changed his/her behavior so the price series deviate from CE and become fluctuating again albeit the tendency towards CE near the end of the *I* treatment. Similarly, in the top-right plot, the buyer changed his/her behavior so the price series deviates from CE and fluctuate between CE and MP in the *I* treatment. In the middle two plots, both the seller and the buyer adjusted their behavior so that the price series become much more volatile in the *I* treatments. It appears that these subjects tried to take advantage of other robot agents by making use of the additional information. On the contrary, there are also cases where subjects became aware of the existence of equilibrium with the help from the new information, and the price series eventually converged to CE or MP in the *I* treatment, as captured by the bottom two plots in Fig. 10.10.

If our subjects reacted to the additional information in different ways, two questions therefore emerged: First, is such information helpful for our human subjects to gain more profit than they did in the *NI* treatment? Secondly, does the introduction of the information improve market efficiency? Table 10.3 compares the means of profits and market efficiency before and after the information revelation. Although it appears that both profits and market efficiency become higher in the last four periods of the *I* treatment, the differences are not statistically significant.[8]

10.4 Conclusion

How prices are formed in various market environment has been an important research agenda in experimental economics. A series of studies have been conducted to investigate whether market price would converge toward competitive equilibrium as suggested by equilibrium theory. Inconsistent results among individual participants and between studies have been found. We aimed at providing a cognitive explanation for the heterogeneous individual behavior, if it ever exists. Our first hypothesis is that subjects with higher cognitive ability, which was measured by CRT, will be in a better position in exploiting their market power and manipulate market prices in favor of their own benefits. On the basis of a series of experimental studies, our sec-

[8] However, this may be because that the market efficiency has been very high in the last four periods of the *NI* treatment. Since there is not much room to improve, the market efficiency cannot be largely improved by the disclosure of information at the first place.

ond hypothesis suggests that markets consisting of participant with higher cognitive ability will converge more quickly to competitive equilibrium.

To address our research questions, we have designed our experiments in a way that can isolate the behavioral influences of every human participant. In this study, each human participant was playing against a group of robot agents, which were designed based on Easley and Ledyard (1993)'s bidding procedures. By doing so, we can discuss the influences of subjects' behavior by comparing the market price series. Our results showed that, counter-intuitively, those who were able to exercise their market power are subjects with lower CRT scores, although the evidence is weak and insignificant. Also, there is no evidence that higher CRT scores were associated with better price convergence. Finally, the revelation of information did alter some subjects' behavior drastically. However, such effects only existed for a few subjects, and the additional information did not increase subjects' performance nor the market efficiency in an economically meaningful manner.

There are limitations to our experimental study. First, the number of subjects is small compared to modern standards. Secondly, the behavior of the Easley-Ledyard bidding agents could be analysed further using large scale simulation so as to facilitate our analyse of aggregate market outcomes due to human subject behavior. Thirdly, CRT is not actually a formal psychometric test, and it doesn't really measure cognitive ability, either. A better candidate would be, for example, working memory capacity. Fourthly, the revelation of information could be studied further if we test more market environment.

Acknowledgements The authors gratefully acknowledged the research support in the form of Taiwan Ministry of Science and Technology grant MOST 106-2410-H-029-009.

Appendix: Instructions

Introduction

Welcome to participate in today's experiment. What you are about to participate in is a "bilateral bidding". Your decisions in the experiment will determine your profits, which will be converted to cash for you. You will be paid privately in cash for your participation at the end of the experiment.

Payoff

Your payoffs will consist of two parts: attendance fee and bonus. Participants in the experiment today will receive NT$100 as an attendance fee. The bonus will be determined based on the profits you earn in the market. The higher the profits you earn, the more bonus you will get after the experiment. Your profits in the experiment will be divided by 20 as your bonus.

Number of Participants in Each Group

In today's experiment, the market will consist of 5 buyers and 5 sellers. You can see your buyer's or seller's identity and serial number from the transaction screen. You will trade with computer traders, which means that all your opponents in the market are computer software, not other participants on the scene. Therefore, please focus on trading activities with computer traders, your profits will not be affected by the decisions of others present. Please keep quiet and do not talk to other participants.

Trading Goal

In the bilateral bidding market, there will be many buyers bidding to buy goods, and many sellers bidding to sell goods. In today's experiment, the buyer's goal is to buy the virtual products from the sellers to make profits; the seller's goal is to sell the products to the buyers to make profits.

If you are a buyer, we will assign you several "tokens" as the basis for the transaction. Please imagine the buyer's tokens as the value of goods to you. When you buy a good, the profit is the difference between the value of the token and the price of the good. The value of the tokens we assign to you may be different. For a buyer, the values of tokens will be ranked from high to low. The buyer must start the transaction with the highest token.

Example 1

A buyer has three tokens, the values of which are 100, 50, and 30 respectively. He must start the transaction with the token value of 100. Suppose he bought the first good at a price of 60, then the profit generated by the first good is $100 - 60 = 40$. Next, he must continue the transaction from the second token. Assuming that he still bought the second good at 60, then the profit generated by the second good is $50 - 60 = -10$. If the third good is bought at a price of 29, what will his profit be? What are the total profits of this buyer in this session?

If you are a seller, the situation is just the opposite of the buyer. We also assign you several tokens as the basis of the transaction. Please imagine the seller's tokens as the costs of goods to you. When you sell a good, your profit is the difference between the price of the good and the value of the token. The values of the tokens we assign to you may be different. For a seller, the value of tokens will be ranked from low to high. The seller must start the transaction with the lowest token.

Example 2

A seller has three tokens, the values of which are 10, 60, and 80 respectively. He must start the transaction with the token value of 10. Suppose he sold the first good at a price of 40, then the profit generated by the first good is $40 - 10 = 30$. Next, he must continue the transaction from the second token. Assuming that he still sold the second good at 40, then the profit generated by the second good is $40 - 60 = -20$. If the third good is sold at a price of 79, what will his profit be? What is the total profits of this seller in this session?

Our experiment today will be conducted for 8 sessions, each session will last 150 s. Each session, every trader in the market will have 5 tokens. Please note that only successful transactions will generate profits or losses. In addition to the profit or loss generated by the transaction, for each successful transaction of a token, regardless of the profitability of the transaction, an additional 5 units of profit can be obtained.

Tokens do not necessarily have to be traded. Untraded tokens will neither accumulate to the next session nor generate profit or loss. We will reissue tokens to you at the beginning of each session. Please note that your goal is to earn the highest total profits possible in exchange for high rewards.

Bargaining Rules and Transaction Prices
In bilateral bidding, the buyer expresses his willingness to buy through bidding, and the seller expresses his willingness to sell through bidding. The bids of buyers and sellers will be recorded in the order book, the bids of buyers will be sorted from high to low, and the bids of sellers will be sorted from low to high. If the buyer's highest bid is greater than or equal to the seller's lowest bid, the bilateral bidders can conclude a transaction, and the transaction will be based on the first bid sent to the order book as the transaction price.

Example 3

Assuming that the current buyer's highest bid is 10 and the seller's lowest bid is 12, there will be no transaction. If a seller bids 9, the buyer's bid is greater than or equal to the seller's bid (10 >= 9), the transaction occurs and the transaction price is 10.

Please enter an integer bidding. If you regret your bid, you can cancel the bid as long as it has not been fulfilled. If you have used up all the tokens before the end of the trading session, you will not be able to bid again during the current session, and your tokens will not be refilled until the next session starts.

References

Bergman O, Ellingsen T, Johannesson M, Svensson C (2010) Anchoring and cognitive ability. Econ Lett 107:66–68

Brañas-Garza P, García-Muñoz T, González RH (2012) Cognitive effort in the beauty contest game. J Econ Behav Organ 83:254–260

Breaban A, Noussair CN (2015) Trader characteristics and fundamental value trajectories in an asset market experiment. J Behav Exp Financ 8:1–17

Burks S, Carpenter J, Goette L, Rustichini A (2009) Cognitive skills affect economic preferences, strategic behavior, and job attachment. Proc Natl Acad Sci 106(19):7745–7750

Calvillo DP, Burgeno JN (2015) Cognitive reflection predicts the acceptance of unfair ultimatum game offers. Judgm Decis Mak 10(4):332–341

Casari M, Ham J, Kagel J (2007) Selection bias, demographic effects, and ability effects in common value auction experiments. Am Econ Rev 97(4):1278–1304

Cason TN, Friedman D (1993) An empirical analysis of price formation in double auction markets. In: Friedman D, Rust J (eds) The double auction market: institutions, theories, and evidence. Westview Press, pp 253–283

Cason TN, Friedman D (1996) Price formation in double auction markets. J Econ Dyn Control 20(8):1307–1337

Cason TN, Williams AW (1990) Competitive equilibrium convergence in a posted-offer market with extreme earnings inequities. J Econ Behav Organ 14(3):331–352

Corgnet B, Hernán-González R, Kujal P, Porter D (2015) The effect of earned versus house money on price bubble formation in experimental asset markets. Rev Financ 19:1455–1488

Corgnet B, Espín AM, Hernán-González R, Kujal P, Rassenti S (2016) To trust, or Not to trust: cognitive reflection in trust games. J Behav Exp Econ 64:20–27
Davis DD, Harrison GW, Williams AW (1993) Convergence to nonstationary competitive equilibria: an experimental analysis. J Econ Behav Organ 22(3):305–326
Davis DD, Williams AW (1986) The effects of rent asymmetries in posted offer markets. J Econ Behav Organ 7(3):303–316
Davis DD, Williams AW (1991) The Hayek hypothesis in experimental auctions: institutional effects and market power. Econ Inq 29(2):261–274
Devetag G, Warglien M (2003) Games and phone numbers: Do short-term memory bounds affect strategic behavior? J Econ Psychol 24:189–202
Devetag G, Warglien M (2008) Playing the wrong game: an experimental analysis of relational complexity and strategic misrepresentation. Games Econom Behav 62:364–382
Duffy J (2006) Agent-based models and human subject experiments. In: Tesfatsion L, Judd KL (eds) Handbook of computational economics, vol 2. Elsevier, pp 949–1011
Easley D, Ledyard JO (1993) Theories of price formation and exchange in oral auctions. In: Friedman D, Rust J (eds) The double auction market: institutions, theories, and evidence. Westview Press, pp 63–97
Fischbacher U (2007) Z-Tree: Zurich toolbox for ready-made economic experiments. Exp Econ 10(2):171–178
Frederick S (2005) Cognitive reflection and decision making. J Econ Perspect 19(4):25–42
Friedman D, Ostroy J (1995) Competitivity in auction markets: an experimental and theoretical investigation. Econ J 105(428):22–53
Gode DK, Sunder S (1993) Allocative efficiency of markets with zero-intelligence traders: market as a partial substitute for individual rationality. J Polit Econ 101(1):119–137
Hayek FA (1945) The use of knowledge in society. Am Econ Rev 35(4):519–530
Hoffman E, Plott CR (1981) The effect of intertemporal speculation on the outcomes in seller posted offer auction markets. Quart J Econ 96(2):223–242
Holt CA (1995) Industrial organization: a survey of laboratory research. In: Kagel JH, Roth AE (eds) The handbook of experimental economics. Princeton University Press, pp 349–443
Holt CA, Langan LW, Villamil AP (1986) Market power in oral double auctions. Econ Inq 24(1):107–123
Hoppe EI, Kusterer DJ (2011) Behavioral biases and cognitive reflection. Econ Lett 110:97–100
Jones G (2008) Are smarter groups more cooperative? Evidence from Prisoner's Dilemma experiments 1959–2003. J Econ Behav Organ 68:489–497
Kagel JH (2004) Double auction markets with stochastic supply and demand schedules: call markets and continuous auction trading mechanisms. In: Huck S (ed) Advances in understanding strategic behavour: game theory. palgrave macmillan, experiments and bounded rationality, pp 181–208
Kahneman D, Frederick S (2002) Representativeness revisited: attribute substitution in intuitive judgment. In: Gilovich T, Griffin D, Kahneman D (eds) Heuristics and biases. Cambridge University Press, pp 49–81
Ketcham J, Smith VL, Williams AW (1984) A comparison of posted offer and double auction pricing institutions. Rev Econ Stud 51(4):595–614
Kiss HJ, Rodriguez-Lara I, Rosa-García A (2016) Think twice before running! bank runs and cognitive abilities. J Behav Exp Econ 64:12–19
Miller RM, Plott CR, Smith VL (1977) Intertemporal competitive equilibrium: an empirical study of speculation. Quart J Econ 91(4):599–624
Noori M (2016) Cognitive reflection as a predictor of susceptibility to behavioral anomalies. Judgm Decis Mak 11(1):114–120
Noussair CN, Tucker S, Xu Y (2016) Futures markets, cognitive ability, and mispricing in experimental asset markets. J Econ Behav Organ 130:166–179
Oechssler J, Roider A, Schmitz PW (2009) Cognitive abilities and behavioral biases. J Econ Behav Organ 72:147–152

Plott CR (1986) Laboratory experiment in economics: the implications of posted-price institutions. Science 232:732–738
Plott CR, Smith VL (1978) An experimental examination of two exchange institutions. Rev Econ Stud 45(1):133–153
Porter DP, Smith VL (1995) Futures contracting and dividend uncertainty in experimental asset markets. J Bus 68:509–541
Roth AE (1995) Introduction to experimental economics. In: Kagel JH, Roth AE (eds) The handbook of experimental economics. Princeton University Press
Schnusenberg O, Gallo A (2011) On cognitive ability and learning in a beauty contest. J Econ Educ 11(1):13–24
Segal N, Hershberger S (1999) Cooperation and competition between twins: findings from a Prisoner's Dilemma game. Evol Hum Behav 20:29–51
Smith A (1937) The wealth of nations. Random House, New York (originally 1776)
Smith VL (1962) An experimental study of competitive market behavior. J Polit Econ 70(2):111–137
Smith VL (1964) Effect of market organization on competitive equilibrium. Quart J Econ 78(2):181–201
Smith VL (1967) Experimental studies of discrimination versus competition in sealed-bid auction markets. J Bus 40(1):56–84
Smith VL (1982) Markets as economizers of information: experimental examination of the Hayek hypothesis. Econ Inq 20(2):165–179
Smith VL, Suchanek GL, Williams AW (1988) Bubbles, crashes, and endogenous expectations in experimental spot asset markets. Econometrica 56:1119–1151
Smith VL, Williams AW (1982) The effects of rent asymmetries in experimental auction markets. J Econ Behav Organ 3(1):99–116
Smith VL, Williams AW (1990) The boundaries of competitive price theory: convergence, expectations, and transaction costs. In: Green L, Kagel J (eds) Advances in behavioral economics, vol 2. Ablex Publishing, Norwood, NJ
Smith VL, Williams AW, Bratton WK, Vannoni MG (1982) Competitive market institutions: double auctions vs sealed bid-offer auctions. Am Econ Rev 72(1):58–77
Stanovich KE, West RF (2000) Individual differences in reasoning: implications for the rationality debate. Behav Brain Sci 23:645–665
Tai C-C, Chen S-H, Yang L-X (2018) Cognitive ability and earnings performance: evidence from double auction market experiments. J Econ Dyn Control 91:409–440
Walker JM, Williams AW (1988) Market behavior in bid, offer, and double auctions. J Econ Behav Organ 9:301–314
Williams FE (1973) The effect of market organization on competitive equilibrium: the multi-unit case. Rev Econ Stud 40(1):97–113
Williams AW (1979) Intertemporal competitive equilibrium. In: Smith VL (ed) Research in experimental economics, vol 1, pp 255–278
Williams AW, Smith VL (1984) Cyclical double-auction markets with and without speculators. J Bus 57(1):1–33

Chapter 11
Counterfactual Thinking and Causal Mediation: An Application to Female Labour Force Participation in India

Sunil Mitra Kumar and Ying-Fang Kao

Abstract The use of computers has revolutionised our ability to learn about ourselves and the world around us. Beyond the goal of performance or prediction, the extent to which machine actions and algorithms are explainable and intelligible to human beings—Explainable AI—are increasingly becoming important, especially so in socio-economic contexts, and where life and health outcomes are involved. While the Turing test aims to distinguish between machine and human, Judea Pearl's 'mini' Turing test focuses on one crucial aspect of this distinction: the ability to reason causally and thereby answer causal queries based on counterfactuals. At the heart of counterfactual-based reasoning lies the role of causal explanation, or delineating the underlying causal mechanisms. In this chapter we make a small step towards demonstrating how causal models can be brought to observational data to answer useful counterfactual queries in contexts where complex social processes are at play. We estimate the causal effects of education on female labour-force participation in India in a causal mediation framework. We consider the role of positive assortative marital-matching in terms of education which leads to husbands' levels of education mediating the effect of women's education on their subsequent labour force participation, and we use a g-formula based approach to estimate the total causal and natural direct effect of education.

Keywords Causal inference · Causal mediation · Female labour force participation

S. M. Kumar (✉)
India Institute and Department of International Development,
King's College London, London, UK
e-mail: sunil.kumar@kcl.ac.uk

Y.-F. Kao
Experimentation Team, Machine Learning and AI Division, Just Eat, London, UK
e-mail: ying-fang.kao@just-eat.com

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_11

11.1 Introduction and Motivation

The use of computers has revolutionised our ability to learn about ourselves and the world around us, and to solve problems, be they academic or practical. A range of approaches are now available to utilise computing machines to learn or infer knowledge from data, adapt and exhibit behaviour (often in complex environments), and together these can be termed *intelligent* in the broad sense of the term. These approaches, which are related in many ways and yet distinct in their aims, can be roughly classified into two broad categories: artificial and machine intelligence. The former involves developing computing machines geared towards understanding or mimicking human intelligence and cognitive capabilities. The latter approach, machine intelligence, is not strictly limited to human capabilities alone, and involves diverse applications across domains that make use of machine learning tools, for e.g. neural networks, classification algorithms, probabilistic learning methods, reinforcement learning and deep learning etc.[1] Although there is considerable debate on these matters, the role of an agent (human or artificial)—how they perceive information, adapt to their environment, decide and act in a goal-oriented manner—retains an important place in both traditions. High-performance, or maximising the chance of achieving a particular goal, often by learning from vast amounts of data, is no longer the absolute criteria to judge such algorithms. Rather, the extent to which these solutions or actions by machines are explainable and intelligible to human beings—Explainable AI—are increasingly becoming important too (Barredo Arrieta et al. 2020; Gunning et al. 2019). This is especially so in socio-economic contexts, and where life and health outcomes are involved.

An important link between the two traditions of AI is the Turing test (Turing 1950) which has played a critical role in creating models or expert systems.[2] The Turing test was originally meant as a way to discern whether the machine or program in question can behave and learn just like human beings in a specific game setting. This can be useful in judging the level of performance, however, questions regarding the process by which such performance is achieved can be important too. While recent efforts in AI focus on building algorithms capable of human or expert level performance, it is still worth bearing in mind that human beings are capable of far deeper learning even with—compared to machines—extremely limited data and computational (cognitive) capacities (Kao and Venkatachalam 2021). For instance, despite the success of deep learning in programs trained to play the game Go, there remains a significant gap between human capability and machine learning. In a similar vein, discussions around different versions of the Turing test acknowledge the gap between human and machine intelligence in socio-economics contexts (Pagliari et al. 2021).

---

[1] These twin categories are sometimes referred to as Classical and Modern AI (Russell and Norvig 2009), or symbolic and connectivist approaches to AI.

[2] These are systems that are capable of yielding valuable insights and performing tasks as required once they have been trained with existing data.

Judea Pearl, recipient of the Turing Award in 2011, proposes an interesting idea to test modern algorithms. Instead of a binary test like what Turing initially considered, Pearl proposes a Ladder of Causation, with progressively more advanced causal queries. Intellectual ability of a program or being can then be graded according to the level of causal reasoning. Pearl terms this the mini Turing Test (Pearl and Mackenzie 2018). With the ability to observe association at the bottom level, and counterfactual thinking at the top, Pearl argues that even the most advanced deep learning algorithms in data science do not go beyond the first rung of the Ladder. Given that counterfactual thinking—an ability to imagine and reason about potential worlds—is ubiquitous in everyday life and causal reasoning is an important arsenal in human intelligence, endowing algorithms with this ability becomes an exciting prospect.

There are several ways in which counterfactual thinking has been invoked in the literature: Chen and Du (2017) explore the link between counterfactual thinking in the context of generalised reinforcement learning and cognitive capacity of human subjects in relation to a beauty contest game. Learning patterns are tied to the notion of counterfactual thinking, since the latter are an important component in evaluating the payoffs associated with strategies under different, imaginary scenarios. Specifically, counterfactual thinking is characterised as in the attraction updating function, called the principle of simulated effect (Camerer and Ho 1998). By including the principle of 'simulated effect', the experience-weighted attraction model generalises reinforcement learning and belief learning and allows for the individual agent to have varying degrees of counterfactual thinking-ability. Chen and Du (2017) are inspired largely by a bulk of psychological literature on associations between the level of counterfactual thinking and working-memory capacity. The primary hypothesis in this literature is that undertaking counterfactual thinking burdens the memory, which implies that heterogeneity in the level of counterfactual thinking could be explained by variations in the level of working-memory capacity. The experimental data from Chen's beauty contest game supports this hypothesis, and overall the study addresses the importance of counterfactual thinking in learning and decision-making in repeated economic games.

However, although not explicitly mentioned in the psychology and cognitive science literature, counterfactual thinking is not feasible without a map of causal pathways or structure. Pearl (2001, 2009) and Imai et al. (2010, 2011) provide a mathematical framework to formalise such structures using Directed Acyclic Graphs (DAGs), and discuss the identification conditions under which various causal, counterfactual-based queries can be answered using observational data. In this chapter our aim is to bring together some of the strands in these literatures, and to make a small step towards demonstrating how causal models can be brought to observational data to answer useful counterfactual queries. We do so in the context of the social sciences and economics. In this realm, answering causal queries remains a complex challenge, and the bulk of efforts focus on estimating treatment effects in contrast to delineating the underlying causal mechanisms. Even if we set aside the practical and ethical limitations involved in the use of randomised controlled trials with human beings in order to obtain unbiased causal estimates (Deaton and Cartwright 2018; Imbens and Rubin 2015), we note that such experiments are better suited to obtaining reduced-form estimates and but however are limited so far as understanding causal mechanisms go. There thus remains significant scope to develop methods that can help understand causal mechanisms.

Our demonstration focuses on an important socio-economic phenomena: the causal effects of education on female labour-force participation. As we discuss, the dominant approach in the literature on this topic focuses on establishing the causal relationship between education and labour force participation (LFP), and thus doing obviates the tougher problem of providing causal *explanation* for this relationship. Our attempt is to shed some light on the latter, and we study this problem in the context of India. Employing insights from the literature on marital matching, we propose a causal structure by which female education effects LFP not only directly, but also indirectly via the husband's level of education. Doing so, we hope to substantiate our claim that methods of causal inference designed to answer counterfactual-based queries in complex causal systems can be productively put to use in the context of socioeconomic phenomena involving human beings and the highly complex social systems within which they exist.

11.1.1 Female Labour Force Participation in India as a Problem of Causal Mediation Analysis

As economies develop, female LFP is believed to initially decrease and later increase, or in other words the association between the two phenomena ought to follow a U-shaped relationship (Klasen 2019), even though empirical evidence suggests that there is significant divergence from this hypothesis at the level of individual countries depending on historical, cultural and economic context (Klasen et al. 2021; Jayachandran 2021).[3] In India, a growing literature focuses on one aspect in particular of economic development: levels of education. There is significant evidence of a neutral-to-negative relationship between rising education and labour force participation amongst married women. Recent studies document that employment rates have remained stagnant (Klasen and Pieters 2015) or declined (Afridi et al. 2018) alongside rising education levels and economic development overall, and it remains a puzzle why this is the case. Afridi et al. (2018) show that this fall has taken place only amongst married women in rural India, and suggest a potential explanation in the form of rising returns to home production relative to labour market participation for educated women. Mehrotra and Parida (2017) document similar trends but provide an alternative explanation. Structural transformation in agriculture has displaced females from this sector, while a combination of rising capital intensity in manufacturing and rising real wages in rural areas has resulted in limited alternative

[3] See Klasen (2019) also for a detailed overview of the challenges and controversies in defining female labour force participation.

employment opportunities, while continuing patriarchal social norms imply limited flexibility all round.

However, average LFP over time is not to be confused with the cross-sectional relationship between LFP and education. As Datta Gupta et al. (2020) find, cross-sectional comparisons of LFP amongst married Indian women as a function of education levels do show a U-shaped relationship. Besides pointing out significant heterogeneity in this overall pattern, their explanations include rising education levels for nonmarket reasons—improved marriage prospects—limited employment opportunities for educated women, and unobserved self-selection processes. Chatterjee et al. (2018) also document a U-shaped relationship, and breaking this down by sector show that rising education correlates with rising salaried employment but declining participation in wage labour, family farms or businesses. As they discuss, in part this points towards the role of limited employment opportunities for educated females, but they also show that the higher household incomes apart from women's own earnings are associated with lower LFP for the women themselves.

The latter point is significant, because marriages in India are still overwhelmingly arranged even as women increasingly participate in spouse-selection with their parents (Allendorf and Pandian 2016). This implies that social norms and specifically the intersection of caste, economic and educational status are still the primary forces that shape marital matches, and could therefore also play a significant role in explaining LFP for married women. As Lin et al. (2020) show, educational hypergamy has decreased in India in line with global trends, reflecting rising female education. However, while educational hypogamy is rising, women who marry down in terms of education are more likely to be marrying into a family with higher economic status than their natal family. The trade-off this evidences, the authors argue, suggests that economic resources play a larger role in marital matching than education levels with the result that women with higher education than their husbands are still likely to earn less than them. This finding is in consonance with Klasen and Pieters (2015)'s list of supply-side factors that depress women's LFP, including higher household incomes and husband's education.

Taken together, these findings have clear implications for better understanding the relationship between female education and LFP. First, that structural factors in the form of limited employment opportunities for educated females, particularly in rural areas likely play an important role. Second, that a combination of societal norms and income substitution effects at household level could be a significant part of the explanation. Third, establishing causality and delineating the underlying causal mechanisms is a consistent challenge throughout this literature. To our knowledge, there is limited evidence on both fronts. The main challenge in inferring causality is that several variables which help explain LFP and are also correlated with education levels are in fact intermediate variables or mediators, for example the education level of husbands and the income of the household as a whole. That is, these are factors causally effected by education and in turn causally effect labour force participation, and this causal structure ought to be explicitly taken into account to provide meaningful insights. Explaining the underlying causal mechanisms is even harder, since there are likely a combination of supply and demand-side factors in operation, and it is challenging to account for all of these simultaneously.

In this paper we address the first of these challenges and make some progress with the second. Specifically, we measure the causal effects of women's education levels on LFP in the cross-section, purposively accounting for the role of social norms within the household as one significant causal explanation that shapes this relationship. To do so we model the assortative matching between women and their husbands in terms of respective education levels, and the causal effects these education levels then have on women's eventual LFP once they are married. Doing so, we attempt to combine some of the insights (Lin et al. 2020) provide for marital matching with their causal implications for the LFP trends documented by Chatterjee et al. (2018). The remainder of the paper is structured as follows. Section 11.2 presents the causal structure we propose to study this relationship, the causal estimands of interest, and estimation methods. Section 11.3 discusses the data used and Sect. 11.4 presents our results. Section 11.5 concludes with a discussion of these results and the new insights offered.

11.2 Causal Structure and Methods

We would expect a woman's level of education to effect her husband's level of education via the process of assortative matching, most likely involving parents on both sides. As Lin et al. (2020) explain, this matching takes into account both education levels but also wealth levels of the respective families. Post-matching, we would then expect the husband's level of education to influence the wife's LFP through pathways including social norms and preferences as well as the husband's earning capacity. The resulting causal structure is portrayed as a Directed Acyclic Graph (DAG) in Fig. 11.1, where husband's education H mediates the effect of female's education E (treatment) on LFP (the outcome).

We focus on two causal estimands. For a given change in female education (E) from level e_0 to e_1, the Total Causal Effect (TCE) measures the resulting change in LFP accounting for both the direct effect of E as well as the indirect effect via the

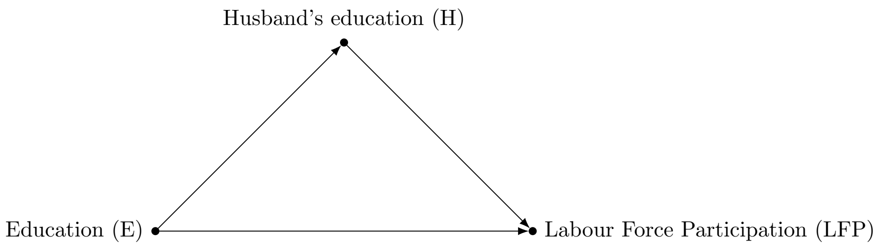


Fig. 11.1 Causal structure for effect of women's education on labour force participation

corresponding change in H. The Natural Direct Effect (NDE) focuses exclusively on one part of this: the change in LFP that is due to the change in E while holding mediator H fixed at the level it naturally attains at the baseline viz. $E = e_0$.[4] These twin causal quantities of interest are defined formally as follows, where $\text{LFP}_{eH_{e*}}$ stands for labour force participation when female education is held at level e and husband's education H is held at the level it would attain when female education is at level e^*.

$$\text{TCE} = E[\text{LFP}_{e_1 H_{e_1}} - \text{LFP}_{e_0 H_{e_0}}] \tag{11.1}$$

$$\text{NDE} = E[\text{LFP}_{e_1 H_{e_0}} - \text{LFP}_{e_0 H_{e_0}}] \tag{11.2}$$

The above definitions apply when LFP is measured as a binary variable. We also consider LFP measured as a 4-category variable that captures the type of work undertaken (none, salaried, farm or business, wage labour). In this case, for each categorical outcome LFP_i where $i \in \{1, 2, 3, 4\}$ these estimands can be defined as:

$$\text{TCE}_i = \text{Prob}[\text{LFP}_{e_1 H_{e_1}} = \text{LFP}_i] - \text{Prob}[\text{LFP}_{e_0 H_{e_0}} = \text{LFP}_i] \tag{11.3}$$

$$\text{NDE}_i = \text{Prob}[\text{LFP}_{e_1 H_{e_0}} = \text{LFP}_i] - \text{Prob}[\text{LFP}_{e_0 H_{e_0}} = \text{LFP}_i] \tag{11.4}$$

11.2.1 Identification and Estimation

Under the overall assumption that the data is generated from a Non-Parametric Structural Equation Model (Pearl 2009, 2014), the following standard assumptions are required for identifying natural direct and natural indirect effects.[5]

A1: **Consistency**: (i) Consistency of E on H. (ii) Consistency of $\{E, H\}$ on LFP. This assumption requires that the actual and potential values coincide for all relevant variables.

A2: **No unobserved mediator-treatment confounding**: There exists a set of variables W_1 such that $H_e \perp\!\!\!\perp E | W_1$. In other words, the effect of treatment on the mediator can be identified by conditioning on W_1.

A3: **No unobserved mediator-outcome confounding**: There exists a set of variables W_2 such that for each $e \in E$ we have that $\text{LFP}_{eH} \perp\!\!\!\perp H | \{W_1, W_2\}$. In other words, holding E fixed, the effect of the mediator H on the outcome can be identified. This also rules out any confounders of the mediator-outcome relationships themselves affected by E.

[4] See Pearl (2001), Imai et al. (2011), Pearl (2014), VanderWeele (2015) for an introduction to causal mediation frameworks more generally.

[5] See VanderWeele (2015, Chap. 2) and Pearl (2014) for a detailed discussion.

A4: **No unobserved treatment-outcome confounding**: There exists a set of variables W_3 such that $\text{LFP}_{eH} \perp\!\!\!\perp E|\{W_2, W_3\}$. In other words, holding H fixed, the effect of treatment on outcome can be identified by conditioning on $\{W_2, W_3\}$. This also rules out any confounders of the mediator-outcome relationship that are themselves affected by the treatment.

Estimating the causal effects listed above requires statistical methods that explicitly model the counterfactuals involved to keep mediators or treatment variables fixed while manipulating the variable of interest, since simply adjusting for the husband's education in a regression of LFP status on women's education and other covariates will in general not yield the required causal estimates.[6] Several methods are available for estimating direct and indirect effects in a causal mediation framework. The simplest of these entails specifying regression models for the mediator and outcome, and using the estimated parameters to provide analytical expressions for the expectation terms in Eqs. 11.1–11.2. This can be challenging particularly when using non-linear regression models since it requires integrating over the conditional counterfactual distribution of the mediator. Monte-Carlo simulation can be used as an alternative, to directly estimate the required expressions through a two-step procedure. In step 1, using the same regression approach of one model each for the mediator and outcome, the estimated model parameters and the covariate values for each observation are then used to generate predictions for the mediator by holding the treatment at required levels. In step 2, using the regression estimates for the outcome model, covariates for each observation having replacing mediators with the predictions, and treatment held at required levels, are then used to generate predicted values for the outcome. Monte-Carlo simulation accounts for the sampling variability of the (conditional counterfactual) distributions of the mediator and outcome. We follow Daniel et al. (2015) and implement this by taking draws from the sampling distribution of the model predictions, obtaining standard errors via bootstrapping (Vansteelandt and Daniel 2017; MacKinnon 2008).

These steps are formalised as follows (see Daniel et al. 2015)

Let $\mathbf{W} = \{W_1, W_2, W_3\}$.

Step 1: Using OLS or maximum likelihood, estimate

$$E(H|E, \mathbf{W}) = g_h(A, \mathbf{W}; \boldsymbol{\beta_h}) \text{ with error variance } \sigma_h^2$$
$$(Y|E, H, \mathbf{W}) = g_y(E, H, \mathbf{W}; \boldsymbol{\beta_y}) \text{ with error variance } \sigma_y^2$$

[6] In this type of regression, the causal interpretation of changing a representative female's education from one level to another in this sort of regression entails averaging over the sample distribution of all other included covariates, whereas the causal quantity of interest requires holding the husband's education fixed. In general, regression does not yield causal estimates if we condition on intermediate outcomes (see Sect. 19.6 in Gelman et al. 2020)—here husband's education—unless we can assume that education does not interact with any other covariate in the model. But the latter is clearly unrealistic since we would expect the causal effect of interest to vary with the husband's level of education. Unfortunately, the simple cure of including an interaction term in the regression specification does not solve the problem either.

Step 2: For each unit i, holding $E = e, e^*$, draw

$$H_i(E, \mathbf{W_i}) \text{ from } N(g_h(E, \mathbf{W_i}, \hat{\boldsymbol{\beta}}_h), \hat{\sigma}_h^2)$$

Step 3: For each unit i, holding $E = e, e^*$ and $E' = e, e^*$, draw $Y_i(E, H_i(E'))$ from $N(g_y(E, H_i(E'), \mathbf{W_i}; \boldsymbol{\beta}_y), \hat{\sigma}_y^2)$.

Step 4: The empirical average across all units i of $Y_i(e^*, H(e))$, $Y_i(e, H(e))$ and $Y_i(e, H(e^*),)$ can be used to estimate the respective expectation terms in (11.1)–(11.2).

11.3 Data and Variables

We use data from the second round of the India Human Development Survey (IHDS).[7] The IHDS is a nationally-representative household survey undertaken in two rounds, 2004–5 and 2011–12 (Desai et al. 2010; Desai and Vanneman 2015), and contains a rich set of questions probing household and individual demographics and labour market participation. A supplementary questionnaire in the 2011–12 survey asks additional questions of a selected ('eligible') woman in a subset of the surveyed households. These questions cover topics including the nature of marital matching, such as the education levels of the woman's own parents and that of her in-laws, and whether the marriage was arranged.

As we now discuss, these covariates are essential for our analysis, and we therefore focus on the 2011–12 data, and within this, women interviewed as part of the 'eligible woman' supplementary questionnaire. We further restrict the analysis to women who have had an arranged marriage and who are in the age group 17–60 and therefore could potentially participate in the labour market. With these restrictions in place, our final sample size is 36,175.

In light of the discussion on causal structure above, an important part of the analysis is deciding which covariates to adjust for. Crucially, we want to avoid adjusting for intermediate outcomes unless they themselves are mediators such as husband's education, that is, unless they causally shape LFP. For this reason, we do not adjust for income or include it in our analysis. First, income is not observable in the dataset at the level of individuals other than for salaried employment. Second and more importantly, if household income is causally determined jointly by the woman's education and the husband's education and goes on to effect LFP, then it is an intermediate outcome and therefore should not be adjusted for. The same applies to household size and more specifically the number of children, since decisions about fertility are also, plausibly, effected by the husband's and wife's education levels jointly, and go on to effect LFP. An additional complication with household size is that a woman might have caring responsibilities for children other than her own in the same household, and thus presents a potential missing data problem.

[7] These are the same data used by Chatterjee et al. (2018).

For these reasons, our analysis uses a smaller number of covariates than the covariates usually adjusted for in the literature. Specifically, our aim is to adjust for the minimal set of covariates that will plausibly fulfill assumptions A1-A4 presented above. In essence, these assumptions refer to identifying three types of causal effects: treatment on mediator, mediator on outcome holding treatment fixed, and treatment on outcome holding the mediator fixed. We now consider each of these in turn.

Identifying the effect of women's own education on that of their husband in fact refers to causal inference in the context of a historic event: the process of marital matching. Given that marriages are arranged, for any given level of the woman's education, her parents' own education (and thus social capital, expectations and aspirations) and income status ought to be sufficient to explain the choice of prospective husband and thus his education. Unfortunately, data on the parents' income are not available, however each woman is asked to list her mother and father's level of education. Of course the nature of marital matching might also vary according to religion and caste, and we therefore control for these as well, and additionally add interaction terms between the parents' education and caste group.

Next, to identify the effect of husband's education on female LFP, we want to adjust for any potential confounders that effect both variables. In the Indian context, this implies adjusting for variables that capture the social attitudes of the husband as well as attitudes of other household members whose views might shape the wife's LFP. To this end, our assumption is that the husband's parents would plausibly have influence over the woman's LFP, and therefore their education levels are a useful proxy for their social attitudes. As above, we recognise that this causal relationship likely varies across caste group and religion, and in the model we estimate below, we therefore include these variables together with their interaction terms with the husband's education. Analogous arguments apply for identifying the effect of the woman's education on her LFP while holding husband's education fixed. That is, beyond religion and caste group, we again assume that the in-laws' (i.e. husband's parents') levels of education are a useful proxy for the attitudes within the household that would help explain the woman's LFP.

Descriptive statistics for the variables used in our analysis are presented in Table 11.1. We additionally also adjust for state (province) fixed effects, however these are not listed in the table for the sake of brevity. The two models described in Step 1 of Sect. 11.2.1 can now be specified as follows

The mediator H (husband's education) is modelled using OLS where the vector $\mathbf{X_h}$ consists of education levels of the woman, her mother and her father, caste and religion dummies, interaction terms between caste-group and the woman's education, a dummy for urban area, and state fixed effects.

$$H = \beta_{h0} + \mathbf{X_h} \beta_\mathbf{h} + e_h \tag{11.5}$$

The outcome LFP is modelled using two separate specifications. For binary LFP we use a logit model where LFP_i represents the labour force participation of the ith female and the vector $\mathbf{X_{LFP}}$ consists of a constant, education and its square, husband's education and its square, education level of the woman's in-laws, caste and religion

Table 11.1 Descriptive statistics

N = 36,175	Mean	S.D.	Min	Max
Labour market participation				
None	0.783	–	–	–
Salaried	0.038	–	–	–
Farm or business	0.061	–	–	–
Wage labour	0.118	–	–	–
Education (years)				
Own education	5.281	(4.933)	0.000	16.000
Mother's education	1.519	(3.079)	0.000	16.000
Father's education	3.509	(4.535)	0.000	16.000
Husband's education	7.127	(4.831)	0.000	16.000
Mother-in-law's education	1.148	(2.697)	0.000	16.000
Father-in-law's education	3.032	(4.263)	0.000	16.000
Age	35.779	(9.631)	17.000	60.000
Urban location	0.336	–	–	–
Religion				
Hindu	0.819	–	–	–
Muslim	0.121	–	–	–
Christian	0.022	–	–	–
Sikh	0.024	–	–	–
Buddhist	0.006	–	–	–
Jain	0.002	–	–	–
Tribal	0.004	–	–	–
Others	0.001	–	–	–
Caste group				
Brahmin	0.052	–	–	–
Forward/General (except Brahmin)	0.236	–	–	–
Other Backward Castes (OBC)	0.406	–	–	–
Scheduled Castes (SC)	0.212	–	–	–
Scheduled Tribes (ST)	0.081	–	–	–
Others	0.000	–	–	–

Notes This table presents means and proportions of all covariates used in the analysis. Survey probability weights are not taken into account in these statistics nor the analysis itself dummies, interaction terms between caste-group and the husband's education as well as the woman's education, a dummy for urban area, and state fixed effects.

$$\text{Prob}[\text{LFP}_i = 1|\mathbf{X}_{\mathbf{i}\text{LFP}}] = \frac{\exp[\mathbf{X}_{\mathbf{i}\text{LFP}}\boldsymbol{\beta}_{\mathbf{LFP}}]}{1 + \exp[\mathbf{X}_{\mathbf{i}\text{LFP}}\boldsymbol{\beta}_{\mathbf{LFP}}]} \tag{11.6}$$

For the categorical version of LFP that takes on four values $\{Y_1, Y_2, Y_3, Y_4\}$, we use a multinomial logit model with the same vector $\mathbf{X_y}$ as before. The vector $\boldsymbol{\beta_l}$ where $l \in \{1, 2, 3, 4\}$ refers to the parameter estimates for outcome Y_l, and $\boldsymbol{\beta_1}$ is set to zero for identification.

$$\text{Prob}[\text{LFP}_i = Y_j|\mathbf{X}_{\mathbf{i}\text{LFP}}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}, \boldsymbol{\beta_4}] = \frac{\exp[\mathbf{X}_{\mathbf{i}\text{LFP}}\boldsymbol{\beta}_j]}{1 + \sum_{l=2}^{4} \exp[\mathbf{X}_{\mathbf{i}\text{LFP}}\boldsymbol{\beta}_l]} \qquad j \in \{1, 2, 3, 4\} \tag{11.7}$$

11.4 Results

Table 11.2 presents the regression estimates for models (11.5), (11.6) and (11.7). The signs of the coefficients show that higher levels of education are associated with higher levels of husband's education, or in other words that there is positive assortative matching on this front, and the positive coefficient for the square of education shows that this relationship is convex. Education also has a convex relationship with LFP overall, but the negative coefficient for education and positive for its square suggest that the exact shape of the LFP-education relationship depends on the specific levels of education being considered. Things are more ambiguous when LFP is considered in terms of types of work and the relationship varies across the categories of work. Higher levels of husband's education however are associated with lower LFP throughout, while the positive coefficient for the square of husband's education indicates a convex relationship for overall (binary) LFP as well as salaried work. Taken together, these estimates suggests that the husband's education is positively effected by the woman's education, and it goes on to effect her LFP negatively, while the woman's own education has a convex association with her LFP that may or may not result in an overall positive association. However, on their own, these coefficients do not allow us to examine the U-shape hypothesis because of the two-step causal structure involved in determining this. Next, we therefore implement the steps provided in Sect. 11.2.1 to estimate the total causal and natural direct effect of education on LFP.

The estimates of causal effects are presented graphically to aid interpretation. Figure 11.2 presents estimated Total and Natural Direct effect based on models (11.5) and (11.6), that is, when the outcome is binary LFP. The base level of education considered is 0 years, i.e. no formal schooling, since this is the mode of the distribution

Table 11.2 Regression models

Dependent variable:	Husband's education	LFP (binary)	LFP by type of work (base: not in work)		
			Salaried	Farm or business	Wage labour
Education	0.379*** (0.022)	–0.084*** (0.025)	0.044 (0.045)	–0.014 (0.040)	0.095 (0.067)
Education × Education	0.002* (0.001)	0.014*** (0.001)	0.014*** (0.001)	0.001 (0.001)	–0.010*** (0.002)
Husband's education		–0.106*** (0.029)	–0.188*** (0.047)	–0.014 (0.045)	–0.102 (0.062)
Husband's edu × Husband's edu		0.002*** (0.001)	0.006*** (0.001)	–0.001 (0.001)	–0.003* (0.001)
Mother's education	0.018* (0.008)				
Father's education	0.128*** (0.006)				
Mother–in–laws's education		–0.019* (0.008)	–0.010 (0.011)	–0.012 (0.014)	–0.098*** (0.018)
Father–in–laws's education		–0.027*** (0.005)	–0.027** (0.008)	–0.024** (0.008)	–0.041*** (0.008)
Caste group (Base category: Brahmin)					
Forward/General (non-Brahmin)	–1.616*** (0.181)	0.263 (0.229)	–0.467 (0.430)	0.353 (0.358)	0.421 (0.443)
Other Backward Castes (OBC)	–2.061*** (0.173)	0.854*** (0.223)	–0.319 (0.411)	0.567 (0.350)	1.119** (0.433)
Scheduled Castes (SC)	–2.670*** (0.177)	1.230*** (0.225)	0.473 (0.413)	0.051 (0.358)	1.827*** (0.434)
Scheduled Tribes (ST)	–3.189*** (0.189)	1.354*** (0.229)	–0.454 (0.461)	1.033** (0.359)	1.849*** (0.437)
Others	–2.067*** (0.319)	0.967*** (0.294)	–0.717 (0.806)	0.704 (0.453)	1.080* (0.493)
Education # Caste group (Base category: Brahmin)					
Forward/General (non-Brahmin)	0.106*** (0.019)	–0.069** (0.023)	–0.030 (0.041)	–0.039 (0.037)	–0.102 (0.066)
Other Backward Castes (OBC)	0.115*** (0.018)	–0.113*** (0.023)	–0.043 (0.039)	–0.056 (0.036)	–0.125 (0.064)
Scheduled Castes (SC)	0.153*** (0.020)	–0.116*** (0.023)	–0.107** (0.040)	–0.050 (0.039)	–0.137* (0.065)
Scheduled Tribes (ST)	0.213*** (0.023)	–0.090*** (0.025)	0.023 (0.048)	–0.078 (0.040)	–0.114 (0.066)
Others	0.131*** (0.038)	–0.092* (0.036)	–0.087 (0.076)	–0.038 (0.057)	–0.045 (0.076)

(continued)

Table 11.2 (continued)

Dependent variable:	Husband's education	LFP (binary)	LFP by type of work (base: not in work)		
			Salaried	Farm or business	Wage labour
Husband's education # Caste group (Base category: Brahmin)					
Forward/General (non-Brahmin)		0.033 (0.027)	0.062 (0.042)	0.006 (0.043)	0.043 (0.062)
Other Backward Castes (OBC)		0.012 (0.027)	0.065 (0.041)	–0.015 (0.041)	0.043 (0.060)
Scheduled Castes (SC)		0.008 (0.027)	0.094* (0.042)	–0.015 (0.043)	0.045 (0.060)
Scheduled Tribes (ST)		0.022 (0.028)	0.072 (0.049)	–0.005 (0.044)	0.053 (0.061)
Others		–0.012 (0.039)	0.173* (0.080)	–0.045 (0.061)	–0.007 (0.072)
Urban	0.609*** (0.044)	–0.536*** (0.035)	0.348*** (0.067)	–0.966*** (0.065)	–0.790*** (0.051)
Constant	6.566*** (0.220)	–1.860*** (0.258)	–3.027*** (0.433)	–3.113*** (0.447)	–4.075*** (0.575)
Religion	Yes	Yes	Yes	Yes	Yes
State (province)	Yes	Yes	Yes	Yes	Yes
N	36321	36300	36300		

Notes This table presents regression estimates for models (11.5), (11.6) and (11.7). Survey probability weights are ignored. Standard errors are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

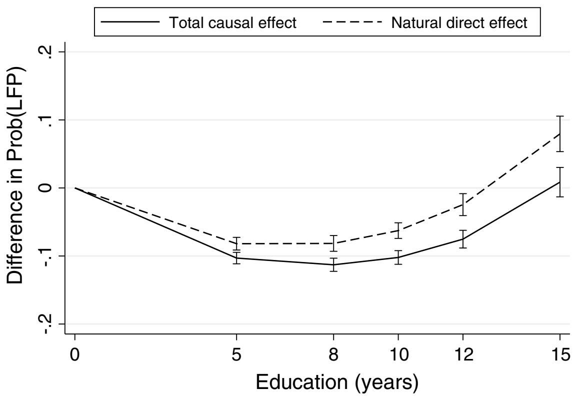


Fig. 11.2 Causal effect of education on female labour force participation (reference level of education = 0)

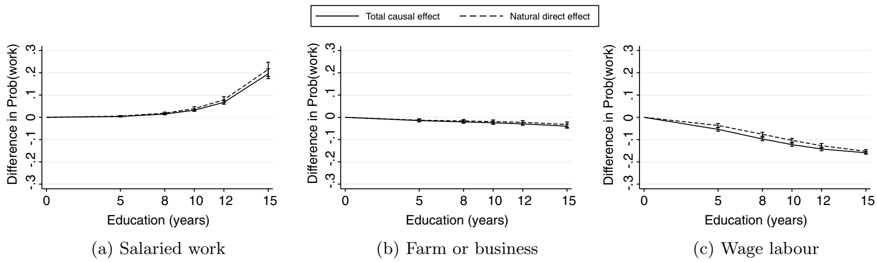
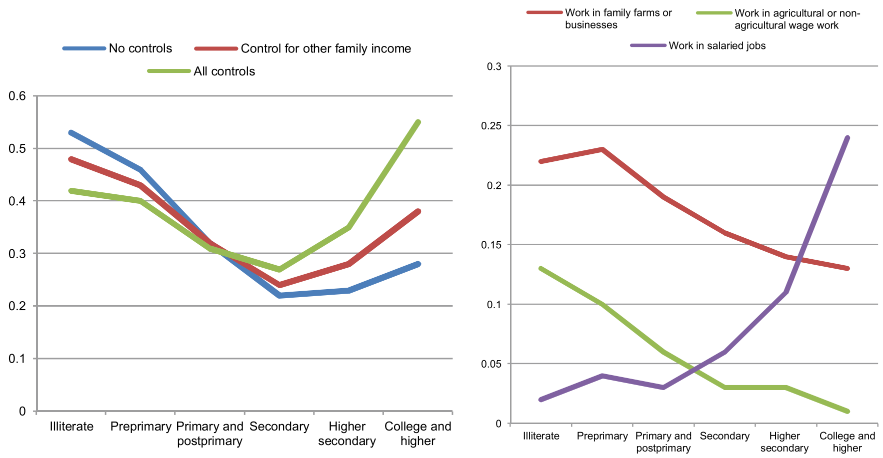


Fig. 11.3 Causal effects of education by type of work (reference level of education = 0)

of education. The x-axis shows the levels of education at which causal effects are estimated. These correspond to the most common levels of education observed in the data which correspond to key stages: primary school (5 years), upper-primary (8), high school (10), higher-secondary (12) and graduate (15). These estimates are based on 200 bootstrap replications for each level of education shown, with 200 Monte-Carlo draws per estimate. Figure 11.2 supports the U-shaped hypothesis: comparing no formal education as the reference level (i.e. education=0) with primary and higher levels, LFP initially decreases, is lowest for upper-primary education, and then begins rising with higher levels of education (post-secondary education in particular). The position of the dotted line (NDE) above the solid line (TCE) shows that the direct causal effect of female education is more positive than its total causal effect. This makes sense, because the latter also takes into account the effect via husbands' education: the positive assortative matching due to marriage, and the negative effect of the husband's education on LFP, both of which factors combine to reduce LFP.

Figure 11.3a–c present the estimates of total and natural direct effects according to type of work. These are based on models (11.5) and (11.7), that is, considering the categorical LFP outcome with three potential types of labour force participation. Other details are the same as those used for the estimates in Fig. 11.2, viz. the base level of education as 0 years, and estimates based on 200 bootstrap replications for each level of education shown, with 200 Monte-Carlo draws per estimate. These figures show that—as we might expect—the results for overall LFP mask significant heterogeneity depending on type of work. Across all three types of work, the TCE and NDE are very similar. For salaried work, the causal effect of education is positive and convex, with LFP rising slightly for primary, upper-primary and senior education, and rising significantly higher for senior-secondary and college education. For farm or business work the relationship is almost a straight line which slopes slightly downwards, suggesting a near-linear, weakly negative relationship, while for wage labour the effects of education are nearly linear and more strongly negative. Recognising that wage labour accounts for the largest proportion of work, the U-shaped relationship for overall LFP therefore reflects the weighted average across these three types of work.

How do these results compare with the literature? The most direct comparison is with Chatterjee et al. (2018) whose main results are reproduced in Fig. 11.4 for ease



(a) Overall LFP (b) By type of work

Fig. 11.4 Chatterjee et al. (2018)'s results: Predicted probabilities of labour force participation by education level. *Source* Chatterjee et al. (2018) p. 869 (overall LFP) and p. 871 (LFP by type of work)

of comparison. To reiterate, our results are based on a far smaller set of covariates for reasons discussed above, and, moreover, our estimation method is based on an account to explicitly model the causal structure involved. With these two significant differences in mind however, the results are very similar. The regression approach of Chatterjee et al. (2018) is most similar to estimating the NDE (albeit under stronger assumptions, including that husband's education is in effect measured pre-treatment). Comparing Fig. 11.4a with Fig. 11.2 shows that if we take the mediator-role of husband's education into account, the total causal effect of education is lower than the direct effect (in regression terms, the 'residual' effect), reflecting the negative role of (positive) assortative marital matching based on education combined with the negative effect of husband's education on women's LFP. To our mind this is a significant advantage of using causal mediation analysis, since it explicitly distinguishes between the two ways in which education effects LFP. Comparing Fig. 11.4b with Fig. 11.3 shows that while the sign and magnitude of effects by type of LFP are similar, our estimates are more attenuated.

11.5 Discussion

Our motivation in this paper is to demonstrate how causal inference and counterfactual reasoning can be applied to disentangle complex social interactions using observational data. While counterfactual-based thinking is a well-studied psychological and cognitive phenomenon (Chen and Du 2017) which has a critical role in advancing modern Artificial Intelligence (Pearl and Mackenzie 2018) and many other domains besides including in the social sciences and economics, there is substantial variation in the attention paid to underlying causal mechanisms. In particular, the toolkits developed for causal mediation analysis in psychology and epidemiology are rarely deployed in economics to analyse social processes. By specifying the question of how education effects women's participation in the labour force in India as a problem of causal mediation, we have attempted to bring these literatures together, and demonstrate how causal mediation analysis can be a valuable tool in this domain.

We have argued why it is important to model education-based assortative marital matching as part of the overall causal explanation for how education effects labour force participation. Doing so explicitly recognises the role of the husband's education as a mediator via which female education shapes labour force participation. Similar to the literature, we find that the overall relationship between education and female labour force participation is U-shaped. However, we do so by considering the total causal effect and the natural direct effect separately. This lends the important insight that while both effects are U-shaped, the total effect is smaller because it also includes the negative effect of the husband's education—where the latter rises with rising female education due to positive assortative marital matching. We also consider the type of labour force participation across three categories, viz. wage labour, salaried work, and farm or business work. The effects of education vary significantly across the three types of work. Total and natural direct effects are very similar, with positive effects for salaried, negative for wage, and very slightly negative for farm or business work.

While the role of the husband's education as a mediator between (female) education and labour force participation is important to recognise, our causal structure nevertheless simplifies matters. For instance, it is plausible that the number of young children in the household also effects labour force participation, and is itself effected by education via fertility decisions and is thus an additional mediator. Further, fertility decisions might also be shaped by the husband's level of education level. Thus the number of young children might be a second mediator itself effected by husband's education as the first mediator. While this is just one example, in general, a more complex causal structure might be called for to study a phenomenon as complex as the labour force participation decision. Through our simple demonstration, our aim therefore is to argue for deploying causal mediation analysis as a valuable tool for investigating complex social processes, where reduced-form causal effects are indeed valuable, but far deeper insights can be gathered by understanding the underlying causal mechanisms.

References

Afridi F, Dinkelman T, Mahajan K (2018) Why are fewer married women joining the work force in rural India? A decomposition analysis over two decades. J Popul Econ 31(3):783–818
Allendorf K, Pandian RK (2016) The decline of arranged marriage? marital change and continuity in India. Popul Develop Rev 42(3):435
Arrieta B et al (2020) Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion 58:82–115
Camerer C, Ho T-H (1998) Experience-weighted attraction learning incoordination games: probability rules. Heterogeneity, and time-variation. J Math Psychol 42:305–326
Chatterjee E, Desai S, Vanneman R (2018) Indian paradox: rising education, declining women's employment. Demogr Res 38:855–878
Chen S-H, Du Y-R (2017) Heterogeneity in generalized reinforcement learning and its relation to cognitive ability. Cogn Syst Res 42:1–22
Daniel RM, Stavola BLD, Cousens SN, Vansteelandt S (2015) Causal mediation analysis with multiple mediators. Biometrics 71(1):1–14
Datta Gupta N, Nandy D, Siddhanta S (2020) "Opt out" or kept out? the effect of stigma, structure, selection, and sector on the labor force participation of married women in India. Rev Develop Econ 24(3):927–948
Deaton A, Cartwright N (2018) Understanding and misunderstanding randomized controlled trials. Soc Sci Med 210:2–21
Desai S, Vanneman R (2015) India human development survey-II (IHDS-II), 2011-12. ICPSR36151-v2. Inter-university Consortium for Political and Social Research [distributor]
Desai S, Vanneman R, National Council of Applied Economic Research, New Delhi (2010) India Human Development Survey (IHDS), 2005. ICPSR22626-v8. Inter-university Consortium for Political and Social Research [distributor]
Gelman A, Hill J, Vehtari A (2020) Regression and other stories. Cambridge University Press, Cambridge
Gunning et al. (1998) (2019) XAI-Explainable artificial intelligence. Sci Robot 4(37):eaay7120
Imai K, Keele L, Tingley D (2010) A general approach to causal mediation analysis. Psychol Methods 15(4):309
Imai K, Keele L, Tingley D, Yamamoto T (2011) Unpacking the black box of causality: learning about causal mechanisms from experimental and observational studies. Am Polit Sci Rev 105(4):765–789
Imbens G, Rubin D (2015) Causal inference: for statistics. Cambridge University Press, Cambridge, Social and biomedical sciences, p 2015
Imbens GW (2020) Potential outcome and directed acyclic graph approaches to causality: relevance for empirical practice in economics. arXiv:1907.07271 [stat]
Jayachandran S (2021) Social norms as a barrier to women's employment in developing countries. IMF Econ Rev 1–20
Kao Y-F, Venkatachalam R (2021) Human and machine learning. Comput Econ 57:889–909
Klasen S (2019) What explains uneven female labor force participation levels and trends in developing countries? World Bank Res Observ 34(2):161–197
Klasen S, Le TTN, Pieters J, Santos Silva M (2021) What drives female labour force participation? comparable micro-level evidence from eight developing and emerging economies. J Develop Stud 57(3):417–442
Klasen S, Pieters J (2015) What explains the stagnation of female labor force participation in urban India? World Bank Econ Rev 29(3):449–478
Lin Z, Desai S, Chen F (2020) The emergence of educational hypogamy in India. Demography 57(4):1215–1240
MacKinnon D (2008) Computer intensive methods for mediation models. Introduct Stat Mediat Anal 325–346

Mehrotra S, Parida JK (2017) Why is the labour force participation of women declining in India? World Dev 98:360–380

Pagliari C, Bucciarelli E, Chen S-H (2021) Challenges in the study of intelligent machines and reverse turing test on socio-economic decisions. Decision economics: minds machine, and their society. Bucciarelli E, Chen S-H, Corchado JM, Parra J (ed). Springer

Pearl J (2001) Direct and indirect effects. In: Proceedings of the seventeenth conference on uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc, pp. 411–420

Pearl J (2009) Causality, 2nd edn. Cambridge University Press, New York

Pearl J (2014) Interpretation and identification of causal mediation. Psychol Methods 19(4):459–481

Pearl J, Mackenzie D (2018) The book of why. Penguin Random House

Russell S, Norvig P (2009) Artificial intelligence: a modern approach, 3rd edn. Prentice-Hall, New York

Turing AM (1950) Computing machinery and intelligence. Mind 59(236):433–460

VanderWeele TJ (2015) Explanation in causal inference: methods for mediation and interaction. Oxford University Press, New York

Vansteelandt S, Daniel RM (2017) Interventional effects for mediation analysis with multiple mediators. Epidemiology (Cambridge, Mass.) 28(2):258–265

Chapter 12
When Supply and Demand do NOT Meet: Sraffa's Critique of Economic Theory Restated

Stefano Zambelli

Abstract The *neoclassical Walrasian microeconomists* assume that the agents are price takers and, concurrently, that these prices are exactly those that would ensure market clearing across all markets. In this chapter, traditional Sraffian schemes are generalized to the case in which there are non-uniform rates of profits and a population of *heterogeneous neoclassical utility maximizing agents*. A series of *computational experiments* are conducted to verify whether this generalisation alters Sraffa's conclusion concerning the indeterminacy of general economic equilibrium. The conclusion is that *market clearing equilibria* in the market for consumption goods *do not occur* for most of the economies we have considered. In other words, *the so-called "law" that supply meets demand does not seem to hold for most cases*.

Keywords Heterogeneous agents · Indeterminacy general equilibrium · Sraffian schemes · Income distribution · Computable economics

12.1 Prelude to a Critique of Economic Theory

In this chapter I restate the major result/critique presented in the book *Production of Commodities by Means of Commodities* (PCMC), Sraffa (1960) using algorithmic methods. Sraffa's book has a very clear and illuminating subtitle: *Prelude to a*

This chapter is a contribution in homage to the work of my colleague and friend Prof. Shu-Heng Chen. We met in 1991 when our mentor Prof. Kumaraswamy Velupillai founded the Center for Computable Economics at the Department of Economics, UCLA. Prof. Shu-Heng Chen and myself share the conviction that algorithmic computation is an important tool for the furthering of economic theory. What I present here is based on digital computations, which I consider to be an appropriate way to honor Prof. Shu-Heng Chen's lifetime research.

S. Zambelli (✉)
Department of Economics and Management, University of Trento, via Inama, 5, 38122 Trento, Italy
e-mail: stefano.zambelli@unitn.it

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_12

Critique of Economic Theory. My claim in this chapter is that Sraffa's book is not just a "prelude to a critique", but a very clear and devastating one.

The essence of the critique is that standard economic theories do not contain enough ingredients to demonstrate that bookkeeping *equilibrium prices*[1] are determined by the working of *free* or *perfect* competition. The implication of this critique, if it is valid, is not only that economic theories are not closed (self-contained), but also, and most importantly, that from these theories we cannot infer that the forces of the markets alone determine distribution (Bharadwaj 1963). In this chapter, I demonstrate that Sraffa's critique is valid.

A very positive side of the critique is that it supports the view that not every attempt to change the distribution is futile, simply because distribution itself is not determined exclusively by the economic forces of *free competition*.[2]

In particular, Sraffa has shown that *values* (i.e., self-replacing prices) and *distribution* cannot be computed independently of each other (Bharadwaj 1963). Therefore, given a distribution, we can calculate the prices, and vice versa. As he pointed out as early as 1925 and 1926 (Sraffa 1925, 1926).

> A STRIKING feature of the present position of economic science is the almost unanimous agreement at which economists have arrived regarding the theory of competitive value, which is inspired by the **fundamental symmetry existing between the forces of demand and those of supply**, and is based upon the assumption that the essential causes determining the price of particular commodities may be **simplified** and grouped together so as to be represented by a **pair of intersecting curves of collective demand and supply**. This state of things is in such marked contrast with the controversies on the theory of value by which political economy was characterised during the past century, that it might almost be thought that from these clashes of thought the spark of an ultimate truth had at length been struck Sraffa (1926, p. 535, *emphasis added*)

What Sraffa showed in 1960 with PCMC, strengthening the arguments already put forward in 1925 and 1926, is that the *theories* of the functioning of the economic system where commodities are produced with other commodities do not contain enough ingredients to allow *the computation of equilibrium prices* through the theoretical construct of a pair of intersecting curves of collective demand and supply.[3]

---

[1] Defined as the uniform prices, profit rates and wage rates that are such that, for each period, the revenues of each agent (whether a producer or a consumer) are exactly equal to their expenditures. This is an equilibrium because it implies that the aggregate supply of goods (revenue) is equal to the aggregate demand (expenditure). This notion of equilibrium pertains to many theoretical settings where demand and supply are assumed to intersect. For reasons of space I cannot elaborate further, but this is a very general notion of equilibrium that pertains to parts of (and in some cases all) the theories developed by the *classical economists* (like Smith, Ricardo, Malthus, Marx, and many others) and to the *marginalist* economists (like Walras, Pareto, Marshall, Edgeworth, Böhm-Bawerk, Wicksell, Hayek, Hicks, Samuelson, Arrow, Debreu, Hahn, Scarf, MasColell, and many others).

[2] On this point see Sraffa Papers, D3/12/111/149, Venkatachalam and Zambelli (2021b, p. 467) and Appendix 12.10.

[3] Although the critique in (Sraffa 1925, 1926) is principally directed against Marshall (1890, Book V. *General Relation of Demand, Supply, and Value*), it also applies to *all* theories where supply and demand are assumed to intersect. The core of Sraffa's reasoning is that when quantities change

What Sraffa had made very clear in 1926 is demonstrated with greater rigor in 1960.

As early as 1926, Sraffa wrote

> …the **assumption** [that the conditions of **production** and the **demand** for a commodity can be considered, in respect to small variations, as being practically **independent**, both in regard to each other and in relation to the supply and demand of all other commodities (p. 538)] becomes **illegitimate**, when a variation in the quantity produced by the industry under consideration **sets up a force which acts directly, not merely upon its own costs, but also upon the costs of other industries**; in such a case the conditions of the "particular equilibrium" which it was intended to isolate are upset, and **it is no longer possible, without contradiction, to neglect collateral effects** Sraffa (1926, p. 539, *emphasis added*)

And in 1960, Sraffa added

> …as the wages fall **the price of the product** of a low proportion [of labor to means of production] …**may rise or it may fall, or it may even alternate in rising and falling, relative to its means of production**; while the price of the product of a high-proportion [of labor to means of production] …industry **may fall, or it may rise or it may alternate**. What neither of such products can do …is **to remain stable in price relative to its means of production** throughout any range, long or short, of the wage-variation (PCMC, pp. 14–15, *emphasis added*).

> …It may be added that not only in this case but *in general* the use of the term **'cost of production'** has been avoided in this work, as well as the term **'capital'** in its quantitative connotation, at the cost of some tiresome circumlocution. This is because these terms have come to be inseparably linked with the supposition that they stand for *quantities that can be measured independently of, and prior to, the determination of the prices of the product*. (Witness the 'real cost' of Marshall and the lqquantity of capital' which is implied in the marginal productivity theory.) Since to achieve **freedom from such presuppositions** has been one of the aims of this work, avoidance of the term seemed the only way of not prejudicing the issue. (Sraffa 1960, p. 9, *emphasis added*)

Sraffian self-replacing prices are *equivalent* to Marshallian, Walrasian, Austrian or any other form of equilibrium prices, in the sense that *the quantity sold* by the producers and workers must have, in order to allow self-replacement, *the same value as the quantities bought* by the producers and workers. The prices that allow this to occur can not be computed independently from the distribution of the surplus.

12.2 Sraffa's Method for His Critique: The Self-replacing System

In 1960 Sraffa was able to address a critique that went to the core of any economic theory based on the notion of *market clearing* equilibrium prices, by studying the

---

in one market or industry, the induced changes in prices would change the cost of production of the other commodities, so that the demand and supply curves would change their positions and *might not intersect*. This is also the case for Walrasian general equilibrium, where all markets are interconnected.

properties of the exchange values (prices) that would be necessary for the system to reproduce itself.

It is a constructive method. If the self-replacing prices could be computed simply with the knowledge of the means of production used in the production cycle (including labor) and of the output produced by each individual industry, the system would be determined. But Sraffa showed that this is not the case. The system is indeterminate.[4]

The different theoretical general **equilibrium** systems have in common the idea of no change. Sraffa, in the Preface (p. v) of PCMC states this very clearly:

> No changes in output …and no changes in the proportions in which different means of production are used by an industry are considered, **so that no question arises as to the variation [of the proportions of the factors of production] or constancy of returns.**
>
> The investigation **is concerned exclusively with such properties of an economic system as do not depend on changes in the scale of production or in the proportions of 'factors'**.
>
> The marginal approach requires attention to be focused on change, **for without change either in the scale of an industry or in the 'proportions of the factors of production' there can be no marginal product nor marginal cost**. In a system in which, day after day, production continues unchanged in those respects, the marginal product of a factor (or alternatively the marginal cost of a product) would not merely be hard to find – it just would not be there to be found[5] (emphasis added)

Sraffa's argumentation is subtle.

Any notion of equilibrium is tautologically related with a system which is not changing. Sraffa, with his device of self-replacing prices, is rigorously exploiting this point. In PCMC he is not discussing at all the assumptions that theorists (like Walras, Marshall, Clark, Jevons, Böhm-Bawerk, Hayek, Hicks and many others) may have made regarding the dimensions of the firms, the properties of the markets and industries, the functioning of the money and credit markets (neutrality of money), the (marginal) utility or (marginal) profits of maximizing agents, and so on, but studies the issue of the *determinants of the equilibrating prices*, once it is assumed that the system has to be reproduced.

The focus is not on the assumptions a theorist might have made with respect to the decisions of the individuals and/or the functioning of markets, but on the *logical implications relative to the computation of equilibrium prices*. In light of the results of PCMC, what a theoretician should not do is to conclude (or assume) that self-

---

[4] His demonstration is based on the observation that the number of equations of the system is less than the number of variables (PMCM, p. 11). The method of counting equations and variables was essential to the founders of general equilibrium theory (Donzelli 2006, p. 494). But "The mere agreement of the number of equations and unknowns is not sufficient to conclude that the equations are soluble" (Wald 1951, p. 403). See also Appendix 12.11.

[5] Many have inferred from this passage that PCMC is a book assuming or dealing with long-run positions. This is not, in my view, correct. Sraffa is stressing here the theoretical notion of equilibrium which he wants to criticize, and does criticize in PCMC. He is not at all implying, as we are going to see below, that the system he deals with is a settled system, but is studying what the equilibrium prices should be, not at all implying that these equilibrium prices have been actual prices or will be actual prices.

replacing prices may be determined exclusively by natural forces independently from distribution and/or power (see Venkatachalam and Zambelli 2021a, b).

If prices could be computed knowing nothing more than the means of production, this would mean that also the distribution among all the individuals composing the economic system (investors and consumers or capitalists and workers) would be mechanically determined. But this is not the case (neither theoretically nor empirically).

12.3 The System to Be Replicated and the Computation of Self-replacing Prices

12.3.1 Period of Production, the Harvest, and the Market Day

Sraffa's book PCMC begins with the following:

> Let us consider an extremely **simple society** …. **Commodities** are **produced** by **separate industries** and are exchanged for one another at a **market** held **after** the **harvest** (PCMC, p. 3, emphasis added).

The principal point in PCMC (at least in Parts 1 and 2) is to determine the

> …**set of exchange-values** which **if adopted** by the market restores the original distribution of the products and **makes it possible for the process to be repeated**; (PCMC, p. 3), emphasis added

so that the system is in a *self-replacing state*.

12.3.2 The Production Cycle and the Harvest Under Self-replacement

The output b_i produced of commodity i, in one production period, has to be re-produced for the next period. The commodities n to be re-produced may be represented by the following vector.

$$[b_1, b_2, \dots b_i, \dots, b_n]^T \tag{12.3.1}$$

$i = 1, 2, \dots, n$.

During the market day, the owners of commodity i have to exchange the quantity b_i against the means of production and labor used in the just concluded production period. This *re-production* may be represented by the sequential relation

$$b_i \xrightarrow{exchange} a_i^1, a_i^2, \dots, a_i^j, \dots a_i^{n-1}, a_i^n, \ell_i \xrightarrow{production} b_i, \tag{12.3.2}$$

where a_i^j is the quantity of means of production j used in the production of the quantity b_i of good i, For the system as a whole, this *circularity* is summarized with the following notation, where the rows are industries:

$$\begin{array}{cccccccccc}
b_1 & \xrightarrow{exchange} & a_1^1 & \ldots & a_1^j & \ldots & a_1^n & \ell_1 & \xrightarrow{production} & b_1 \\
b_2 & \xrightarrow{exchange} & a_2^1 & \ldots & a_2^j & \ldots & a_2^n & \ell_2 & \xrightarrow{production} & b_2 \\
\vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \xrightarrow{\cdots\cdots\cdots} & \vdots \\
b_i & \xrightarrow{exchange} & a_i^1 & \ldots & a_i^j & \ldots & a_i^n & \ell_i & \xrightarrow{production} & b_i \\
\vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \xrightarrow{\cdots\cdots\cdots} & \vdots \\
b_n & \xrightarrow{exchange} & a_n^1 & \ldots & a_n^j & \ldots & a_n^n & \ell_n & \xrightarrow{production} & b_n
\end{array} \tag{12.3.3}$$

These sequences of harvest–exchange–production may be written in compact matrix notation in the following way:

$$\mathbf{b} \xrightarrow{exchange} \mathbf{A}, \boldsymbol{\ell} \xrightarrow{production} \mathbf{b} \tag{12.3.4}$$

where: $\mathbf{A}$ is an $n \times n$ matrix whose components are the used means of production $\{a_i^j\}$; $\boldsymbol{\ell}$ is a $n \times 1$ vector whose elements $\{\ell_i\}$ is the labor used in production; $\mathbf{b}$ is $n \times 1$ vector whose element $\{b_i\}$ is the harvest of good i.[6]

12.3.3 The Surplus Available for Distribution and Its Consumption

Note that the quantities of Eq. 12.3.3 or Eq. 12.3.4 also include the surplus that was produced during the previous production period, and that will be replicated also in the subsequent one. That is:

[6] In the rest of the chapter we will use bold lowercase letters for vectors, with the exception of the labor vector $\mathbf{L}$ and the physical surplus vector $\mathbf{S}$. Uppercase letters will be used for matrices. To simplify the notation, row i of a matrix, let us take as example the matrix $\mathbf{A}$, will be written in bold in the following way, $\mathbf{a}_i$, while column j will be written as $\mathbf{a}^j$. Furthermore we will assume all the values to be rational numbers. This is a question of realism (quantities and prices in actual economic systems are always observed to be rational numbers with a finite number of digits) and unavoidable, given the algorithmic nature of the theoretical investigation conducted in this chapter. Addition, subtraction, multiplication, and division of rational numbers can in principle be made with absolute precision. Also, the inverse of a matrix whose elements are rational numbers can be made with absolute precision (Aberth, 2007, Chap. 9).

$$
\begin{aligned}
s_1 &= b_1 - \sum_{i=1}^{n} a_i^1 \\
s_2 &= b_2 - \sum_{i=1}^{n} a_i^2 \\
\vdots &= \vdots - \vdots \\
s_j &= b_j - \sum_{i=1}^{n} a_i^j \\
\vdots &= \vdots - \vdots \\
s_n &= b_n - \sum_{i=1}^{n} a_i^n
\end{aligned}
\tag{12.3.5}
$$

where s_i is the surplus of commodity i available for distribution after the quantities $\{a_i^j\}$ have been put aside for next year's production, or, alternatively, is the quantity produced in the previous period which is left once the inputs used in production have been removed.

In compact matrix notation, we have:

$$\mathbf{s} = (\mathbf{B} - \mathbf{A})^T \mathbf{e} \tag{12.3.6}$$

where: $\mathbf{e}$ is an $n \times 1$ unit or summation vector (each element is 1); T is the transpose operator; $\mathbf{s}$ is the $n \times 1$ surplus vector.

12.3.3.1 Consumption

We are considering a system that reproduces itself, therefore, the surplus vector $\mathbf{s}$ is also the *social consumption vector* $\mathbf{c}$. Clearly the distribution of the surplus among the different agents participating in the production process is in terms of the consumption of each agent. It could not be otherwise.

We can list the agents (both producers and workers) with the index $\omega = 1, 2, \ldots, \omega, \ldots, n_\omega$, where n_ω is the number of agents in the system.

Agent ω consumes

$$\mathbf{c}^\omega = \begin{bmatrix} c_1^\omega \\ \vdots \\ c_i^\omega \\ \vdots \\ c_n^\omega \end{bmatrix} \tag{12.3.7}$$

where: $\mathbf{c}^\omega$ is a column vector of dimension $n \times 1$, and c_i^ω is the consumption by agent ω of commodity i.

The above individual consumption vectors may be organized in the following *social consumption matrix*:

$$\mathbf{C} = \begin{bmatrix} c_1^1 & c_1^2 & \dots & c_1^\omega & \dots & c_1^{n_\omega} \\ c_2^1 & c_2^2 & \dots & c_2^\omega & \dots & c_2^{n_\omega} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ c_i^1 & c_i^2 & \dots & c_i^\omega & \dots & c_i^{n_\omega} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ c_n^1 & c_n^2 & \dots & c_n^\omega & \dots & c_n^{n_\omega} \end{bmatrix}_{n \times n_\omega} = \left[\mathbf{c}^1, \mathbf{c}^2, \dots, \mathbf{c}^\omega, \dots, \mathbf{c}^{n_\omega}\right]_{n \times n_\omega}; \tag{12.3.8}$$

Clearly, each column is the consumption vector associated to one agent.

The sum by row of **C** is the *aggregate social consumption vector* **c** or surplus **s**. The following must hold:

$$\mathbf{c}_{n \times 1} = \mathbf{s} = \mathbf{C}\,\mathbf{e}_{n_\omega \times 1} \tag{12.3.9}$$

e is an $n_\omega \times 1$ unit or summation vector. Clearly, the sum by columns of **C** does not have any economic meaning, because it would be the sum of non-homogeneous quantities.

The population may be separated into two groups: Producers, whose number is n_π, and workers, whose number is n_ℓ.[7] The total population is obviously the sum of the two groups: $n_\omega = n_\pi + n_\ell$.

Therefore, we can partition **C** into the matrix containing all the information regarding the consumption of the producers, ${}^{n_\pi}\mathbf{C}_{(n+1) \times n_\pi}$, and another matrix describing all the consumption of the workers, ${}^{n_\ell}\mathbf{C}_{(n+1) \times n_\ell}$.

That is:

$$\mathbf{C} = \left[\,{}^{n_\pi}\mathbf{C} \;:\; {}^{n_\ell}\mathbf{C}\,\right] \tag{12.3.10}$$

Consequently, the *aggregate social consumption vector* **c** may be separated into consumption by the producers, $\mathbf{c}^\pi$, and consumption by the workers, $\mathbf{c}^\ell$. That is:

$$\mathbf{c} = \mathbf{c}^\pi + \mathbf{c}^\ell = [\,{}^{n_\pi}\mathbf{C}\,]\,\mathbf{e}_{n_\pi \times 1} + [\,{}^{n_\ell}\mathbf{C}\,]\,\mathbf{e}_{n_\ell \times 1} \tag{12.3.11}$$

The matrix **C** and vectors **c**, $\mathbf{c}^\pi + \mathbf{c}^\ell$ are all quantities of commodities.

The expenditure of each agent in consumption is given by the following

$$\mathbf{c}^{Expenditures} = \mathbf{C}^T\mathbf{p} \tag{12.3.12}$$

which is a vector of dimension $n_\omega \times 1$ and **p** is the vector of prices which has dimension $n \times 1$.

Equation 12.3.12 may be divided into the value of personal consumption expenditures of the producers and that of the workers:

[7] This is a classification done for convenience. Some producers could also have labor as a source of income, and some workers could also be owners of firms and also have as income the payment of dividends.

$$\mathbf{c}^{Expenditures}_{n_\omega \times 1} = \begin{bmatrix} \mathbf{c}^{Expenditures\ Producers} \\ \mathbf{c}^{Expenditures\ Consumers} \end{bmatrix} = \begin{bmatrix} {}^{n_\pi}\mathbf{C}^T \\ \hline {}^{n_\ell}\mathbf{C}^T \end{bmatrix} \mathbf{p} \tag{12.3.13}$$

12.3.4 Definition of Self-replacing Prices

The *self-replacing* or *market clearing* prices are the prices for which the following accounting balances hold (Zambelli 2018b):

$$\begin{aligned} (1+r_1)\mathbf{a}_1\mathbf{p} + \ell_1 w &= b_1 p_1 \\ (1+r_2)\mathbf{a}_2\mathbf{p} + \ell_2 w &= b_2 p_2 \\ \ldots &= \ldots \\ (1+r_i)\mathbf{a}_i\mathbf{p} + \ell_i w &= b_i p_i \\ \ldots &= \ldots \\ (1+r_n)\mathbf{a}_n\mathbf{p} + \ell_n w &= b_n p_n \end{aligned} \tag{12.3.14}$$

where: $\mathbf{p}$ is the vector of uniform prices; $\mathbf{a}_1, \mathbf{a}_2 \ldots \mathbf{a}_i \ldots \mathbf{a}_n$ are $1 \times n$ row vectors, of the means used in production; $r_1, r_2, \ldots, r_i \ldots r_n$ are positive industry *profit rates*; w is the *wage rate*.

In matrix notation, Eqs. 12.3.14 become

$$(\mathbf{I} + \mathbf{R})\mathbf{A}\mathbf{p} + \boldsymbol{\ell} w = \mathbf{B}\mathbf{p} \tag{12.3.15}$$

where: $\mathbf{R} = diag(\mathbf{r})$ is a diagonal matrix whose diagonal elements are the profit rates of each single industry, $r_1, r_2, \ldots, r_i, \ldots r_n$ (vector $\mathbf{r}$).[8]

12.3.5 Distribution of the Surplus

12.3.5.1 Functional Distribution of Income

The *distribution of the surplus* to the different industries and to the workers is determined by the self-replacing prices and wage rate. We have:

[8] The accounting implicit in Eqs. 12.3.14 and 12.3.15 is made consistent with the choice made in PCMC. We think that the alternative of computing the profit rates as including also labor costs would be simpler and more appropriate. Nevertheless, the qualitative conclusions would not change. The difference is that, for example, Eq. 12.3.14 would have to be written as $(\mathbf{I} + \mathbf{R})(\mathbf{A}\mathbf{p} + \boldsymbol{\ell} w) = \mathbf{B}\mathbf{p}$. The proposition proved in the sequel of the article may be appropriately modified to consider these different accounting prices.

$$\mathbf{ds}^T\mathbf{p} = \overbrace{\begin{bmatrix} d_1\mathbf{s}^T\mathbf{p} \\ d_2\mathbf{s}^T\mathbf{p} \\ \vdots \\ d_n\mathbf{s}^T\mathbf{p} \\ d_w\mathbf{s}^T\mathbf{p} \end{bmatrix}}^{\substack{\textit{Functional distribution} \\ \textit{of the surplus}}} = \overbrace{\begin{bmatrix} b_1p_1 - \mathbf{a}_1\mathbf{p} - \ell_1 w \\ b_2p_2 - \mathbf{a}_2\mathbf{p} - \ell_2 w \\ \vdots \\ b_np_n - \mathbf{a}_n\mathbf{p} - \ell_n w \\ \mathbf{e}^T\boldsymbol{\ell}w \end{bmatrix}}^{\substack{\textit{Purchasing Capacity} \\ \textit{Income}}} = \overbrace{\begin{bmatrix} r_1\mathbf{a}_1\mathbf{p} \\ r_2\mathbf{a}_2\mathbf{p} \\ \vdots \\ r_n\mathbf{a}_n\mathbf{p} \\ \mathbf{e}^T\boldsymbol{\ell}w \end{bmatrix}}^{\substack{\textit{Purchasing Capacity} \\ \textit{Income}}} \tag{12.3.16}$$

where $d_1\mathbf{s}^T\mathbf{p}, d_2\mathbf{s}^T\mathbf{p}, \ldots, d_n\mathbf{S}^T\mathbf{p}$ are the shares of the Net National Product going to the n industries and $d_w\mathbf{S}^T\mathbf{p}$, is the share going to labor. By definition—or construction—the sum of the elements of the vector $\mathbf{d}$ is 1.

In more compact matrix notation,

$$\mathbf{ds}^T\mathbf{p} = \begin{bmatrix} (\mathbf{B}-\mathbf{A})\mathbf{p} - \boldsymbol{\ell}w \\ \mathbf{e}^T\boldsymbol{\ell}w \end{bmatrix} = \begin{bmatrix} \mathbf{RAp} \\ \mathbf{e}^T\boldsymbol{\ell}w \end{bmatrix} \tag{12.3.17}$$

where $\mathbf{d}$ is the *functional distribution of income* because it is the remuneration of the industries producing the means of production and labor. Here, $\mathbf{R}$ is a diagonal matrix whose diagonal is the vector of profit rates $\mathbf{r}$.

Below we are going to use $\mathbf{s}^T$ as a *numéraire* for the prices which by definition gives $\mathbf{s}^T\mathbf{p} = 1$. This being the case, the term $\mathbf{s}^T\mathbf{p}$ may be removed from the left hand side of Eqs. 12.3.16 and 12.3.17.

12.3.5.2 Personal Distribution of Income

An economic system is composed of agents. The *value added* during the *production cycle* is the net national income to be distributed to the members of the society. The consumption of each agent is represented by the columns of matrix $\mathbf{C}$. The income of each individual agent is, in the case in which the agent is an owner of a firm belonging to an industry, the dividends of the firm's profits in addition to the rewards for the agent's work effort.

The whole *personal income distribution* may be described by the matrix $\mathbf{H}_{(n+1)\times n_\omega}$ (multiplied by $\mathbf{s}^T\mathbf{p}$).

$$\mathbf{Hs}^T\mathbf{p} = \begin{bmatrix} h_1^1 & h_1^2 & \ldots & h_1^\omega & \ldots & h_1^{n_\omega} \\ h_2^1 & h_2^2 & \ldots & h_2^\omega & \ldots & h_2^{n_\omega} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ h_i^1 & h_i^2 & \ldots & h_i^\omega & \ldots & h_i^{n_\omega} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ h_n^1 & h_n^2 & \ldots & h_n^\omega & \ldots & h_n^{n_\omega} \\ h_w^1 & h_w^2 & \ldots & h_w^\omega & \ldots & h_w^{n_\omega} \end{bmatrix} \mathbf{s}^T\mathbf{p} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_i \\ \vdots \\ \mathbf{h}_n \\ \mathbf{h}_w \end{bmatrix} \mathbf{s}^T\mathbf{p} = \tag{12.3.18}$$

$$= \left[\mathbf{h}^1, \mathbf{h}^2, \ldots, \mathbf{h}^\omega, \ldots, \mathbf{h}^{n_\omega}\right]\mathbf{s}^T\mathbf{p};$$

Columns are associated with the persons in the population ($\omega = 1, 2, \ldots, n_\omega$, where n_ω is the total population). The rows are associated with the sources of income, i.e., the n industries-commodities and labor (see also Appendix 12.12).

12.3.5.3 Distribution of the Surplus to the Individual Producers/Firms

In the sequel, in order to simplify the analysis presented in this chapter, we make the assumption that the producers owning a firm belonging to industry i have as a source of their income their share of the total profits made in the industry, but do not have income from other sources. Also, the workers will have their income exclusively from labor. We simplify the analysis also by assuming: that there is only one owner per firm; that the number of owners is also the number of firms; that a producer can own only one firm. Given this simplification, we have that the number of producers belonging to industry 1 is n_{π_1}, to industry 2 n_{π_2}, ..., to industry i $n_{\pi_i}, \ldots,$ to industry n, n_{π_n}.[9]

Therefore $n_\pi = n_{\pi_1} + n_{\pi_2} + \cdots + n_{\pi_i} + \cdots + n_{\pi_n}$. We can partition matrix $\mathbf{H}$ into $n+1$ rectangular matrices. Each row of these matrices is a different source of income (a firm's distributed profits or labor income), while the columns are associated with the members of the producers-firms that belong to that particular industry. The last partition is relative to the workers.

$$\begin{aligned}\mathbf{H}_{(n+1)\times n_\omega} &= [{}^{n_\pi}\mathbf{H} : {}^{n_\ell}\mathbf{H}] = \\ &= [{}^{n_{\pi 1}}\mathbf{H}_{(n+1)\times n_{\pi_1}} : {}^{n_{\pi 2}}\mathbf{H}_{(n+1)\times n_{\pi_2}} : \cdots : {}^{n_{\pi_i}}\mathbf{H}_{(n+1)\times n_{\pi_i}} : \cdots : {}^{n_{\pi n}}\mathbf{H}_{(n+1)\times n_{\pi_n}} : {}^{n_\ell}\mathbf{H}_{(n+1)\times n_\ell}]\end{aligned} \tag{12.3.19}$$

The matrix ${}^{n_{\pi_i}}\mathbf{H}_{(n+1)\times n_{\pi_i}}$, with $i = 1, 2, 3, \ldots, n$, has dimension $(n+1) \times n_{\pi_i}$. It has 0s in all rows, except for the i^{th} row, which has the dividends distributed by industry i. The elements of the i^{th} row are the share of the profits of industry i associated to the firms belonging to the industry, $j = 1, 2, \ldots, n_{\pi_i}$ are the producers belonging to industry i, ${}^{n_{\pi_i}}\mathbf{H}(i, j)$ is the share of the surplus profits of firm j belonging to industry i. Aggregating by rows matrices i, ${}^{n_{\pi_i}}\mathbf{H}(i, j)$ we obtain the following:

$$\mathbf{h} = \mathbf{h}^{n_\pi} + \mathbf{h}^{n_\ell} = \mathbf{h}^{n_{\pi_1}} + \mathbf{h}^{n_{\pi_2}} + \cdots + \mathbf{h}^{n_{\pi_i}} + \cdots + \mathbf{h}^{n_{\pi_n}} + \mathbf{h}^{n_\ell} \tag{12.3.20}$$

or, which is the same,

$$\mathbf{d} = \mathbf{d}^{n_\pi} + \mathbf{d}^{n_\ell} = \mathbf{d}^{n_{\pi_1}} + \mathbf{d}^{n_{\pi_2}} + \cdots + \mathbf{d}^{n_{\pi_i}} + \cdots + \mathbf{d}^{n_{\pi_n}} + \mathbf{d}^{n_\ell}. \tag{12.3.21}$$

Similarly, we can partition the consumption matrix $\mathbf{C}$ (Eq. 12.3.10) in the same way:

[9] Due to limited space, the simplification is made for the sake of the exposition, but it is not a crucial one. For future work, the more realistic case where agents are both owners of firms and workers should be considered. The structure embedded in Eq. 12.3.18 is sufficiently general.

$$\mathbf{C}_{n\times n_\omega} = \left[{}^{n_\pi}\mathbf{C} \;:\; {}^{n_\ell}\mathbf{C}\right] = \\ = \left[{}^{n_{\pi_1}}\mathbf{C}_{n\times n_{\pi_1}} \;:\; {}^{n_{\pi_2}}\mathbf{C}_{n\times n_{\pi_2}} \;:\cdots:\; {}^{n_{\pi_i}}\mathbf{C}_{n\times n_{\pi_i}} \;:\cdots:\; {}^{n_{\pi_n}}\mathbf{C}_{n\times n_{\pi_n}} \;:\; {}^{n_\ell}\mathbf{C}_{n\times n_\ell}\right] \tag{12.3.22}$$

Aggregating by rows, we obtain

$$\mathbf{c} = \mathbf{c}^{n_\pi} + \mathbf{c}^{n_\ell} = \mathbf{c}^{n_{\pi_1}} + \mathbf{c}^{n_{\pi_2}} + \cdots + \mathbf{c}^{n_{\pi_i}} + \cdots + \mathbf{c}^{n_{\pi_n}} + \mathbf{c}^{n_\ell}. \tag{12.3.23}$$

12.4 Computation of Self-replacing Prices

Once the *numéraire*, $\boldsymbol{\eta}$ (whose property is $\boldsymbol{\eta}^T\mathbf{p} = 1$), is picked to be the surplus $\mathbf{s}$, we have by this particular choice of *numéraire* that $\mathbf{s}^T\mathbf{p} = 1$.

Therefore, Eq. 12.3.17 is simplified into the (Sraffa's) fundamental distribution equation[10]:

$$\mathbf{d} = \left[\begin{array}{c|c} (\mathbf{B}-\mathbf{A}) & -\boldsymbol{\ell} \\ \hline \mathbf{0}_{1\times n} & \mathbf{e}^T\boldsymbol{\ell} \end{array}\right]\left[\frac{\mathbf{p}}{w}\right] = \left[\begin{array}{c} \mathbf{RAp} \\ \mathbf{e}^T\boldsymbol{\ell}w \end{array}\right] \tag{12.4.1}$$

There is a unique relation between the price vector $\left[\frac{\mathbf{p}}{w}\right]$ (or the profit rates $\mathbf{R}$ and the wage rate w)[11] and the functional distribution $\mathbf{d}$.

The elements of the share vector $\mathbf{d}$ are bounded from below (value 0) and from above (value 1) and the sum of all the elements is 1. As long as $\mathbf{d}$ has the above properties, the computation of the prices unique and straightforward.

---

[10] Many readers, who are familiar with the interpretation made by many Sraffians and neo-Ricardians of Sraffa's contribution of PCMC may think that this equation is not faithful to Sraffa because it has been generalized to case of non-uniform rates of profits (of which the case of a uniform rate of profits is a specialization). Above, in this chapter, and in Zambelli (2018b), I have shown that Sraffa's choice of the uniform rate of profits is important for the critique of economic theory, but cannot be considered Sraffa's positive view on how the economic system actually works. The equation is a formal description of Sraffa's major result, i.e., that the system is indeterminate. It is an extension of Sraffa's equation of PCMC to the cases where the profit rates are not uniform. This is a faithful description of Sraffa's framework and it is for this reason that I consider it to be Sraffa's "fundamental distribution equation." Readers who are convinced of the importance of the assumption of uniform rates of profits may consider the restriction of the above equation to the case where the diagonal values of $\mathbf{R}$ are all equal. This would be a less general approach, but the major results would not change.

[11] Appendix 12.13 shows how to uniquely determine prices once the vector of profit rates, $\mathbf{r}$ is given.

$$\mathbf{d} = \left[\begin{array}{c|c} (\mathbf{RA}) & \mathbf{0}_{n\times 1} \\ \hline \mathbf{0}_{1\times n} & \mathbf{e}^T\boldsymbol{\ell} \end{array}\right]\left[\frac{\mathbf{p_r}}{w_\mathbf{r}}\right] = \left[\begin{array}{c} \mathbf{RA}\left[\mathbf{B}-(\mathbf{I}+\mathbf{R})\mathbf{A}\right]^{-\mathbf{1}}\mathbf{L}[\boldsymbol{\eta}^\mathbf{T}(\mathbf{B}-(\mathbf{I}+\mathbf{R})\mathbf{A}))^{-\mathbf{1}}\mathbf{L}]^{-1} \\ \mathbf{e}^T\boldsymbol{\ell}[\boldsymbol{\eta}^\mathbf{T}(\mathbf{B}-(\mathbf{I}+\mathbf{R})\mathbf{A}))^{-\mathbf{1}}\mathbf{L}]^{-1} \end{array}\right] \tag{12.4.1r}$$

The price vector $\mathbf{p_r}$ and the wage rate $w_\mathbf{r}$ are computed with the knowledge of the profit rate vector $\mathbf{r}$ (see eqs. 12.13.10 and 12.13.11). Recall that here the *numéraire* is the surplus, i.e., $\eta = \mathbf{s} = (\mathbf{B}-\mathbf{A})^T\mathbf{e}$.

Associated to a feasible functional distribution vector there is a unique vector of self-replacing prices and wage rate, which we might call
(Sraffa's) fundamental self-replacing prices equation:

$$\begin{bmatrix} \mathbf{p}_{\bar{\mathbf{d}}} \\ w_{\bar{\mathbf{d}}} \end{bmatrix} = \left[\begin{array}{c|c} (\mathbf{B}-\mathbf{A}) & -\boldsymbol{\ell} \\ \hline \mathbf{0}_{1\times n} & \mathbf{e}^T\boldsymbol{\ell} \end{array}\right]^{-1} \bar{\mathbf{d}} = \left[\begin{array}{c|c} (\mathbf{B}-\mathbf{A})^{-1} & (\mathbf{B}-\mathbf{A})^{-1}\frac{\boldsymbol{\ell}}{\mathbf{e}^T\boldsymbol{\ell}} \\ \hline \mathbf{0}_{1\times n} & \frac{1}{\mathbf{e}^T\boldsymbol{\ell}} \end{array}\right] \bar{\mathbf{d}} \tag{12.4.2}$$

The subscript $_{\bar{\mathbf{d}}}$ is added to indicate that the prices and the wage rate are computed with the knowledge of the distribution $\bar{\mathbf{d}}$.[12] Obviously this does not imply causality. Once we have determined the prices coherent with a given distribution, we can determine the profit rates associated with that distribution.

We can compute the domain of all possible self-replacing prices and wage rates by computing the values associated with all possible (discrete) feasible distribution vectors $\mathbf{d}$. Since our domain is given in terms of (finite-digit) rational numbers, what we mean here by the phrase "all possible (discrete) combinations of the distribution vector $\mathbf{d}$" is in the sense of a desired precision factor: given any desired level of precision, we will have a finite number of vectors $\mathbf{d}$.[13] This set of vectors is an enumerable set. All possible values can be computed, associated to each of the vectors belonging to the set. For example, if the desired or observed precision is given by m, the number of distributional vectors is an integer which is a function of m and $n+1$. For example, in the case of three commodities and one-digit precision ($m = 0.1$), the number of distributional vectors is 275; with two-digit precision ($m = 0.01$), it is 176,219, and with three-digit precision ($m = 0.001$), it is 167,605,051, and so on.

In short, the set of all possible distributions $\{\mathbf{d}\}$ can be computed with high numerical precision.

From the computational point of view, we have that for any given distribution $\bar{\mathbf{d}} \in \{\mathbf{d}\}$, by implementing Eqs. 12.4.2 and 12.4.3, we can compute the triple $\left(\mathbf{p}_{\bar{\mathbf{d}}}, \mathbf{r}_{\bar{\mathbf{d}}}, w_{\bar{\mathbf{d}}}\right)$ associated to it.

The set of all *self-replacing prices* is fully determined.[14]

$$\left\{\mathbf{d}^q\right\}_{q=1}^{n_d} \quad \longrightarrow \quad \left\{\mathbf{p}^q_{\{\mathbf{d}\}}, \mathbf{r}^q_{\{\mathbf{d}\}}, w^q_{\{\mathbf{d}\}}\right\}_{q=1}^{n_\mathbf{d}} \tag{12.4.4}$$

where n_d is the number of distribution functions, which is determined by the precision factor m and the number of produced commodities (i.e., industries).

From the economics point of view, it is the one to one mapping between distributions and self-replacing prices which is relevant.[15]

[12] Once $\bar{\mathbf{d}}$ is given from Eq. 12.4.2, we have $\mathbf{p}_{\{\bar{\mathbf{d}}\}}$ and by considering Eq. 12.4.1 and dividing the left-hand side element by element we obtain the profit rate $\mathbf{r}_{\{\bar{\mathbf{d}}\}}$. That is,

$$\mathbf{r}_{\bar{\mathbf{d}}} = \bar{\mathbf{d}}_{n\times 1} \oslash \left(\mathbf{A}\mathbf{p}_{\bar{\mathbf{d}}}\right); \quad \mathbf{R}_{\bar{\mathbf{d}}} = diag\left(\mathbf{r}_{\bar{\mathbf{d}}}\right) \tag{12.4.3}$$

.

[13] This method of providing exact procedures is in our view consistent with Sraffa's constructive approach, see Velupillai (1989).

[14] The set $\{\mathbf{p}, \mathbf{r}, w\}$ is decidable.

$$\left\{\mathbf{p}^q, \mathbf{r}^q, w^q\right\}_{q=1}^{n_\mathbf{d}} \quad \Longleftrightarrow \quad \left\{\mathbf{d}\right\}_{q=1}^{n_\mathbf{d}} \tag{12.4.5}$$

Sraffa's major critique is that the knowledge of the methods of production alone is not sufficient to compute the unique triple $(\bar{\mathbf{p}}, \bar{\mathbf{r}}, \bar{w})$ and therefore the distribution of the surplus $\mathbf{s}$ is not determined exclusively by market forces (for example, by "free competition").

Sraffa's critique, in Parts I and II of PCMC, is articulated by picking as independent (exogenous) distribution variable the uniform rate of profit $r \in [0, \Re]$, where $\Re$ is the maximum rate of profits (PCMC, p. 19).

By picking by assumption r to be uniform, Sraffa makes his prelude to a critique of economic theory more evident (Zambelli 2018b). But this does not change the fact that self-replacing prices are not determined.[16]

12.5 Sraffian Schemes and *General Equilibrium*

We can now elaborate on the connection between Sraffian schemes and general equilibrium by implementing a constructive/computational method.

Given the triple $\left(\mathbf{A}, \boldsymbol{\ell}, \mathbf{b}\right)$, we can consider it to be a general equilibrium if certain conditions are fulfilled.

The following is a list of conditions which can be considered as being associated with the notion of general equilibrium:

(a) **Market clearing for produced commodities**. The computed prices, wage rate and profit rates must be such that the following equations are satisfies:

$$\overbrace{(\mathbf{I} + \mathbf{R}_{\{\mathbf{d}\}})\mathbf{A}\mathbf{p}_{\{\mathbf{d}\}} + \boldsymbol{\ell} w_{\{\mathbf{d}\}}}^{\text{Expenditures per industry}} = \overbrace{\mathbf{B}\mathbf{p}_{\{\mathbf{d}\}}}^{\text{Revenues per industry}} \tag{12.3.15}$$

It is not the case that for any triple $\left(\mathbf{p}, \mathbf{r}, w\right)$ the above accounting equality holds. If it does not hold, this means that there will be unsold quantities. Hence the computed prices, wage rate, and profit rates are not equilibrium prices.

Here, equality and market clearing of produced goods is assured because in Sect. 12.4 we have shown how to compute with precision the set of self-replacing prices $\left\{\mathbf{p}_{\{\mathbf{d}\}}, \mathbf{R}_{\{\mathbf{d}\}}, w_{\{\mathbf{d}\}}\right\}$.

The above system of equations has n equations and $2n + 1$ unknowns: n profits rates r_i; n prices p_i and a wage rate, w. The relevant problem for eco-

[15] Bharadwaj (1963, p. 1450)'s review of PCMC is very clear on this point. See also Zambelli (2018b, p. 794) or Venkatachalam and Zambelli (2021b, p. 456).

[16] From the computational point of view, we still have that

$$\left\{r \in [0, \Re]\right\} \Rightarrow \left\{\mathbf{r}\right\}_{Uniform} \Rightarrow \left\{\mathbf{p}, w\right\}_{Uniform} \Rightarrow \left\{\mathbf{d}\right\}_{Uniform} \tag{12.4.6}$$

The problem of what it is that determines the actual value of $r \in \left\{\mathbf{r}\right\}_{Uniform}$ remains.

nomic theory is to determine among the set of all possible accounting equilibria $\{\mathbf{p}_{\{\mathbf{d}\}}, \mathbf{R}_{\{\mathbf{d}\}}, w_{\{\mathbf{d}\}}\}$ which one is, if any, the feasible economic solution. Knowledge of the methods of production alone is not sufficient to determine the feasible solution. This is the essence of Sraffa's critique towards economic theory.

(b) **Market Clearing for consumption goods**. The dividends and the wage incomes of the individual agents (captured by matrix $\mathbf{H}$) are spent on consumption goods in accordance with the following *consumer budget constraints*:

$$\overbrace{\mathbf{y}^{Incomes} = \mathbf{H}^T_{\{\mathbf{d}\}} \left[\frac{\mathbf{p}_{\{\mathbf{d}\}}}{w_{\{\mathbf{d}\}}} \right]}^{Revenues\ per\ consumer} \qquad (12.C.2)$$

$$\overbrace{\mathbf{c}^{Expenditures} = \mathbf{C}^T \mathbf{p}_{\{\mathbf{d}\}}}^{Expenditures\ per\ consumer} \qquad (12.3.12)$$

$$\mathbf{y}^{Incomes} = \mathbf{c}^{Expenditures} \qquad (consumer\ budget\ constraints)$$

(c) **Neutrality of Money**. The market clearing conditions (a) and (b) imply that deferred means of payments are used for the exchanges during the period, but do not constitute a **store of value**. In equilibrium, the values of the commodities bought by the buyers match exactly the values of the commodities sold by the sellers—in this case there is no need for deferred means of payments to be kept as a store of value. If they were kept as a store of value, the *eventual* market clearing equilibrium would be momentary and the market clearing prices would not be self-replacing prices. On the possibility of having temporary market clearing prices with the use of deferred means of payments with the fulfilment of the self-replacing condition, see Venkatachalam and Zambelli (2021a, b)[17];

(d) **Neoclassical general equilibrium market clearing**. There are additional behavioral requirements for the demand for consumption goods. Neoclassical market clearing is defined in terms of optimization (often, the maximization of an objective function). All agents belonging to the system optimize, under given budget constraints, their objective functions.[18]

---

[17] Sraffa (1932) had criticized Hayek, claiming that *money*, and any other financial instrument, being not only a means of exchange, but principally a means for deferred payments (a store of value) is relevant for the determination of the eventual general equilibrium. This was one of the major critiques directed by Sraffa at Hayek (1931) *Prices and Production*: when the economic system moves out of a general equilibrium position, a return to the previous equilibrium requires the use of money in the form of credit and debt relations, which constitutes a store of value, but this would imply *new* equilibrium conditions. Conceptually, once inflation generates conditions of disequilibrium—following Hayek's starting point and reasoning—money (financial magnitudes in terms of debt–credit relations) would begin to generate a situation where the economic system (as depicted by Hayek) would not and could not (according to Sraffa) return to the previous general equilibrium position.
On the exchange between Hayek and Sraffa see also Bracci (2022).

In particular, one has to add to the condition b) above, the condition that $\mathbf{C}$ is such that each individual consumer ω maximizes a well-behaved utility function $U^{\omega}(\mathbf{c}^{\omega})$[19] subject to a budget constraint. That is,

$$\max_{\mathbf{c}^{\omega}} U^{\omega}(\mathbf{c}^{\omega}) s.t. \quad \mathbf{c}^{\omega T}\mathbf{p} = y_{\omega}^{Income} \tag{12.5.1}$$

$\mathbf{c}^{*\omega}$ solves the above neoclassical decision problem. It is the desired or maximizing consumption bundle of agent ω.
The columns of $\mathbf{C}^{*}$ are the vectors of the optimal consumption decision for each consumer.

$$\overbrace{\mathbf{y}^{Incomes} = \mathbf{H}_{\{\mathbf{d}\}}^{T} \begin{bmatrix} \mathbf{p}_{\{\mathbf{d}\}} \\ w_{\{\mathbf{d}\}} \end{bmatrix}}^{Revenue\ for\ each\ consumer} \qquad (12.C.2)$$

$$\overbrace{\mathbf{c}^{*Expenditures} = \mathbf{C}^{*T}\mathbf{p}_{\{\mathbf{d}\}}}^{Expenditure\ for\ each\ consumer} \tag{12.5.2}$$

$$\mathbf{y}^{Incomes} = \mathbf{c}^{*Expenditures} \qquad (consumer\ budget\ constraints) \tag{12.5.3}$$

Each consumer is able to maximize their utility function when there is **no excess demand** (see Sect. 12.7.6.2 and, for example, Arrow and Hahn (1971, Chap. 2, pp. 16–33)).

(e) **Neoclassical factors of production**. An additional requirement for marginalists is that the price of each means of production must be equal to its **marginal productivity**, and the wage rate must be equal to the marginal productivity of labor. This adds additional equations to the system. The self-replacing prices are those that allow the triple $(\mathbf{A}, \boldsymbol{\ell}, \mathbf{b})$ to be replicated in time for the next production cycle.
The output $\mathbf{b}$ could well have been produced by well behaved production functions, by using the factors of production $\mathbf{A}$ and labor $\boldsymbol{\ell}$.
One may assume[20] a neoclassical production function:

---

[18] Typically a producer's objective is the maximization of profits (or minimization of costs) and the consumer's objective is to maximize utility. Both producers and consumers are subject to budget constraints.

[19] "Well-behaved" means $U^{\omega}(\mathbf{c}^{\omega})$ with $\frac{\partial U^{\omega}(\mathbf{c}^{\omega})}{\partial \mathbf{c}^{\omega}} \geq 0$ and $\frac{\partial^2 U^{\omega}(\mathbf{c}^{\omega})}{\partial \mathbf{c}^{\omega 2}} \leq 0$.

[20] And Sraffa allows the reader to do so, see the preface to PCMC. There are no empirical or theoretical grounds that justify the use of such neoclassical production functions, see Zambelli (2018a).

$$b_i = f_i(\mathbf{a}_i, \ell_i) = f_i(a_i^1, a_i^2, \ldots, a_i^j, \ldots, a_i^n, \ell_i)$$
$$\frac{\partial f_i}{\partial a_i^j}(\mathbf{a}_i, \ell_i) \geq 0, \quad \frac{\partial^2 f_i}{\partial a_{ij}^2}(\mathbf{a}_i, \ell_i) \leq 0, \quad \forall i = 1, 2, \ldots, n \text{ and } \forall j = 1, 2, \ldots, n$$
$$\frac{\partial f_i}{\partial, \ell_i}(\mathbf{a}_i, \ell_i) \geq 0, \quad \frac{\partial^2 f_i}{\partial, \ell_i^2}(\mathbf{a}_i, \ell_i) \leq 0, \quad \forall i = 1, 2, \ldots, n \tag{12.5.4}$$

Marginalists would then assume that the prices and the wage rate are equal to the corresponding marginal productivities.[21] That is, for each individual price p_j, it must be the case that

$$p_j = \frac{\partial f_1}{\partial a_1^j}(\mathbf{a}_1, \ell_1) = \frac{\partial f_2}{\partial a_2^j}(\mathbf{a}_2, \ell_2) = \cdots = \frac{\partial f_i}{\partial a_i^j}(\mathbf{a}_i, \ell_i) = \cdots = \frac{\partial f_n}{\partial a_n^j}(\mathbf{a}_n, \ell_n) \tag{12.5.5}$$

and for the wage rate w, it must be the case that

$$w = \frac{\partial f_1}{\partial \ell_1} = \frac{\partial f_2}{\partial \ell_2} = \cdots = \frac{\partial f_i}{\partial \ell_i} = \cdots = \frac{\partial f_n}{\partial \ell_n}. \tag{12.5.6}$$

Conditions (a), (b), and (c) define a general equilibrium, without the neoclassical optimization conditions of (d) and (e).

A first look of the above conditions (e) and (d) may give the idea that once the associated functions are made explicit, the system is now determinate. Condition (e) does provide a specific structure to compute the self-replacing prices and the wage rate, while condition (d) does introduce the utility functions that do formalize the demand side.

But in this case, the logical reasoning or starting point for the search for the equilibrium has to be reversed, going from condition (e) to (d) and then on to (c), (b), and finally (a).

The full set of production functions given by Eqs. 12.5.4 must employ the factors of production in such a way that the output vector be equal to **b**. Given the substitutability between factors, we now have that the matrix of the factors of production **A** is itself a function of prices. It is not guaranteed that **for any set** of well behaved production functions condition (e) is fulfilled. The accounting equilibrium of condition (a) must be fulfilled in all circumstances and this means the production factors of matrix **A** and the labor inputs $\boldsymbol{\ell}$ must be such that Eqs. 12.5.5 and 12.5.6 allow a determination of the prices and of the wage rate.

But these prices and wage rate may not belong to the set of self-replacing prices $\left\{\mathbf{p}_{\{\mathbf{d}\}}, \mathbf{R}_{\{\mathbf{d}\}}, w_{\{\mathbf{d}\}}\right\}$ and may not be such that condition d) is realized.

My **conjecture** is that the indeterminacy remains. The degree of indeterminacy may even increase, due to the strong impositions of conditions (d) and (e). This point

[21] The maximization of profits may take place within different market conditions—monopolies or oligopolies or also when markets are imperfect. The prices of the factors of productions may well be above their marginal productivities, but this is not the standard neoclassical *perfect competition* market clearing assumption.

certainly requires further investigation. For now, the work done by Mandler (1999), although using a different research method from the one applied here, provides some theoretical evidence that Sraffian schemes are, when *set inside a neoclassical Walrasian general equilibrium framework*, indeterminate. Sraffa's self-replacing prices are bookkeeping equilibrium prices that are also Walrasian general equilibrium prices, hence if Sraffa's system is indeterminate, this, in accordance with Mandler (1999)'s result, is also.

Sraffa's critique is very general because it applies to all the theoretical systems where supply equals demand for all markets.

Points (a), (b), and (c), when fulfilled, do imply that supply equals demand for all markets. These conditions do characterize *accounting* market clearing conditions in general. But there is no implication that the agents *optimize* their objective function or that their objective functions are well-behaved neoclassical objective functions.

Clearly, condition (d) is an additional *neoclassical* requirement. If Sraffian schemes are indeterminate, the maximization principle (d) is irrelevant because it can be checked or realized only when prices are known. But indeterminacy means that at least one of these prices is not determined by the so-called "law" that supply equals demand.

12.6 Sraffa's Critique Is Circumvented by Postulation

The neoclassical Walrasian microeconomists assume that the agents are price takers and, concurrently, that the prices are exactly the prices that would ensure market clearing in all markets. If this was not the case, there would be consumers who would not be able—for the given prices—to maximize their utility, and producers who would not be able to maximize profits.[22]

The idea is that competitive market forces would lead to market clearing prices, but this convergence is not formalized in a computationally meaningful way.[23]

[22] Microeconomics textbooks present the consumer's maximization of utility and the firm's maximization of profits in the first chapters with the assumption that both consumers and producers can not influence the prices (and hence they are price-takers) because they are small (with respect to the size of the market). At the same time, the existence of consumers and producers who are not able to trade at the given prices is not at all considered (or is simply ignored). The implication of this is that either exchanges take place with the issuing of deferred means of payments (IOUs or money) or the prices are equilibrium self-replacing prices. The only other possibility is that market clearing takes place but not all consumers or producers maximize their objective functions. This means that conditions (a), (b), and (c) are satisfied, but not (d) and (e). And this is not a neoclassical equilibrium.

[23] The standard mathematical proof of the existence of Walrasian general equilibrium is based on the *existence* of a fixed point (or on the validity of fixed-point theorems). Prominent microeconomists do admit that without proving the existence of a fixed point, general equilibrium can not be shown to exist (Debreu, 1987, p. 218) or cannot be computed (Scarf, 1987, p. 560). On the uncomputability-undecidibility of general equilibrium, see Velupillai (2006), Velupillai (2010, Part II, Chaps. 5, 6, and 7).

Hence, with the neoclassical postulate that there has already been convergence towards the market-clearing prices, the agents are assumed to trade with the other agents as if they did not participate in the convergence process and behave as if they were price-takers all along. Prices are given by some external and mysterious superior force.

Clearly, if the prices, $\mathbf{p}^{neoclassical}$, and wage rate, $w^{neoclassical}$, are given, the functional distribution vector is unique and determined by the externally given prices (see Eq. 12.4.1).

That is:

$$\mathbf{d}^{neoclassical} = \left[\begin{array}{c|c} (\mathbf{B}-\mathbf{A}) & -\boldsymbol{\ell} \\ \hline \mathbf{0}_{1\times n} & \mathbf{e}^T\boldsymbol{\ell} \end{array}\right]\left[\begin{array}{c} \mathbf{p}^{neoclassical} \\ \hline w^{neoclassical} \end{array}\right] \qquad (\textit{Neoclassical distribution})$$

One could be tempted to say that this is a misrepresentation because the $\mathbf{p}^{neoclassical}$ prices are equal to the marginal productivities of the factors of production and the wage rate, $w^{neoclassical}$, is equal to the marginal productivity of labor (in all industries). But this would mean reversing the causal relationship. In fact, in the factor market prices equilibrium, condition (e) (eqs. 12.5.5 and 12.5.6) is assumed to take place at given prices. One is totally free to provisionally postulate that production of output **b** (i.e., **B**) occurs by firms having access to a neoclassical production function (see Eq. 12.5.4) and which maximize profits as in Eqs. 12.5.5 and 12.5.6. But if the system is in equilibrium, the methods used during the production cycle that determined the social output are not the focus, which is to find the prices that would allow replication, i.e., the persistence of equilibrium.

This way of circumventing Sraffa's critique that the system is indeterminate, by *postulating* that prices are fixed at the values where supply equals demand, simply confirms that the forces that render a system, which would be otherwise indeterminate, determinate are not studied, but are simply assumed to exist.

Clearly, this is not an answer to Sraffa's devastating critique. A *solution* to the indeterminacy problem requires that the prices be endogenous. Otherwise, it is not a solution. And Sraffa's critique is not addressed at all. It is simply avoided by postulation.

12.7 An Algorithmic Digital Laboratory for an In-Depth Study

We have set up a an algorithmic digital laboratory with the purpose of highlighting some of the properties or implications of the reasoning developed above.

Table 12.1 Simple economic system

	Means of Production				Output	Surplus
	(1) Iron ↓ $\mathbf{a}^1$	(2) Coal ↓ $\mathbf{a}^2$	(3) Wheat ↓ $\mathbf{a}^3$	(4) Labour[a] $\boldsymbol{\ell}$	(5) **b**	(6) **s**, **c**
Iron Industry $\mathbf{a}_1 \rightarrow$	90	120	60	$\frac{3}{16}$ (45)	180	0
Coal Industry $\mathbf{a}_2 \rightarrow$	50	125	150	$\frac{5}{16}$ (75)	450	165 c_{coal}
Wheat Industry $\mathbf{a}_3 \rightarrow$	40	40	200	$\frac{8}{16}$ (120)	480	70 c_{wheat}
Used, total	180	285	410	1 (240)		

The values in the table are taken from (Sraffa, 1960, p. 19). Some computations using the same numbers may be found in Zambelli (2018b) and in Venkatachalam and Zambelli (2021a)
[a]The total labor force is normalized to 1. The labor force is n_ℓ and its numerical value is defined below (Sect. 12.7.4) as 240. The employment by each industry is in parentheses

12.7.1 Numerical Definition of the Quantities Involved: $(\mathbf{A}, \boldsymbol{\ell}, \mathbf{b})$

We define the triple $(\mathbf{A}, \boldsymbol{\ell}, \mathbf{b})$ with the same numbers used by Sraffa (PCMC, p. 19). Given our focus on PCMC as a *Prelude to a Critique of Economic Theory*, we consider this triple to be the *equilibrium quantities for the Walrasian system* or the equilibrium quantities for the neo-Ricardian gravitation or long-period system. Given the self-replacing condition that there is no net investment, the final consumption vector **c** is the surplus, **s** (see Eqs. 12.3.5, 12.3.6, and 12.3.9).

In this example, there are three goods produced by separate industries: *iron*, *coal* and *wheat*. During the *production period*, before the *harvest*, each industry has used for the production of its output, the *means of production* (in physical quantities) in the proportions of Table 12.1.

The table provides numerical content to the quantities described in Sect. 12.3.2. The matrix of the means of production **A** is composed of the first three columns. The vector of labor inputs $\boldsymbol{\ell}$ is composed of the elements of column 4. The gross output **b** is column 5 and the produced surplus **s** is column 6.

Given the self-replacing condition, there is no net investment, so the final consumption vector **c** is the surplus, **s** (see Eqs. 12.3.5, 12.3.6, and 12.3.9).

Therefore, the owners-firms present in the three industries enter into the market having at their disposal for exchange the output produced and the workers enter into the market with their labor. In this system, the only capital to be used to buy other commodities is the output, owned by the producers, and the labor owned by the workers.

12.7.2 The Set {d} of all Possible Discrete Functional Distributions

For the purpose of this numerical example, we define the set of all possible functional distributions $\{\mathbf{d}\}_{q=1}^{176219}$ as a set with precision $h = 0.01$. This means that there are 176,219 (see Sect. 12.4). That is:

$$\{\mathbf{d}^q\}_{q=1}^{176219} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.99 \\ 0.01 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.99 \\ 0 \\ 0.01 \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0.01 \\ 0 \\ 0.99 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0.01 \\ 0.99 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{12.7.1}$$

12.7.3 Computation of the Domain $\{\mathbf{p}_{\{\mathbf{d}\}}, \mathbf{r}_{\{\mathbf{d}\}}, w_{\{\mathbf{d}\}}\}$ for the Self-Replacing Prices, Profit Rates, and Wage Rates

Given the triple $(\mathbf{A}, \boldsymbol{\ell}, \mathbf{b})$ and the domain of the functional distribution $\{\mathbf{d}\}$, we can compute the domain of the self-replacing prices, profit rates and wage rate $\{\mathbf{p}_{\{\mathbf{d}\}}, \mathbf{r}_{\{\mathbf{d}\}}, w_{\{\mathbf{d}\}}\}$ by implementing Eq. 12.4.2, Sraffa's fundamental self-replacing prices equation.

12.7.4 Population of Producers and Workers: n_π and n_ℓ

As explained in Sect. 12.3.5.3, we separate the population of producers n_π from the population of workers n_ℓ.

For the simulations below, the number of producers is $n_\pi = 60$ and the number of workers is $n_\ell = 240$. In order to study the problem of the distribution of firms over the different industries, we will allow the number of producers-firms per industry to vary as the functional distribution of income varies.

Producers and workers are all consumers, therefore the total number of consumers is the population, $n_\omega = n_\pi + n_\ell$.

12.7.4.1 Firms' or Producer Densities per Industry

With the aim of making our simulations robust, we carried out a set of simulations for varying densities of firms per industry, and hence also varying numbers of producers per industry. In our numerical simulation, the total number of firms is $n_\pi = 60$. Hence there are 1711 possible combinations.

$$\{\mathbf{FirmsDensity}^z\}_{z=1}^{1711} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 58 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 57 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 56 \end{bmatrix}, \ldots, \begin{bmatrix} 57 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 57 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 58 \\ 1 \\ 1 \end{bmatrix} \right\} \tag{12.7.2}$$

The first element of the vectors is the number of firms in the iron industry, the second is for the coal industry, and the third, the wheat industry.

There should be at least one firm per industry. Note that the cases in which there is only one firm in an industry correspond to the case of a monopoly.

12.7.5 Consumer Incomes, **y**

As we have seen in Sect. 12.3.5.2, once we know the number of producers/firms in each industry and the number of workers, we can compute the vector of personal incomes, using the values of the self-replacing prices, profit rates, and the wage rate.

The set of all personal incomes associated with a specific population distribution z is given by

$$\{^{\mathbf{z}}\mathbf{y}^q\}_{q=1}^{176219} \tag{12.7.3}$$

From Eqs. 12.12.6 and 12.12.7 and the information about the population we can derive the average personal income of the workers:

$$y^{\textit{Per capita worker income}} = \frac{\mathbf{e}^T \boldsymbol{\ell} w}{n_\ell} = \frac{\sum_i^{n_\ell} y_i^{\textit{Worker income}}}{n_\ell} = \frac{d_{n+1}}{n_\ell} = \frac{d_w}{n_\ell} \tag{12.7.4}$$

where $d_n + 1 = d_w$ is the workers' share of the net national product. Due to the fact that labor is homogeneous, workers' per capita income is also the income of each worker.

The average or per capita producer's income is given by

$$y^{\textit{Per capita producer income}} = \frac{\mathbf{e}^T \mathbf{RAp}}{n_\pi} = \frac{\sum_i^{n_\pi} y_i^{\textit{Producer incomes}}}{n_\pi} = \frac{\sum_i^{n} d_i}{n_\pi} = \frac{1 - d_{n+1}}{n_\pi} \tag{12.7.5}$$

The producers are not homogeneous and hence an actual individual producer's income is not always equal to the per capita producer income.

12.7.6 Consumption c and C

Following Sraffa's method, here we focus on the condition that would allow the system to replicate itself. There is no physical net investment, and therefore the total consumption vector **c** is the same as the surplus produced by the system (see Sect. 12.3.3.1).

Hence, we have by construction that the market for the means of production and the market for consumption goods both clear for each of the bookkeeping self-replacing prices $\{\mathbf{p}^q\}_{q=1}^{176219}$.

Supply and demand meet perfectly for each of the $\mathbf{p}^q$ self-replacing prices.[24]

Therefore conditions (a) and (b) are fulfilled by construction. But Sraffa's *Prelude to a Critique of Economic Theory* requires that the focus be placed on the notion of **neoclassical** free competition market clearing, where the consumers are defined as utility maximizers.

12.7.6.1 Consumers Defined as Neoclassical Utility Maximizers

As we have described above, when we have the ingredients $\left(\mathbf{d}^q, \mathbf{p}^q, \mathbf{r}^q, w^q, \mathbf{FirmsDensity}^z\right)$, we can compute the associated income distribution vector ${}^{\mathbf{z}}\mathbf{y}^q$.

In order to acquire their desired consumption goods, consumers have to exchange a fraction of their incomes with other consumers. This is the essence of the exchange economy as in the standard microeconomic analysis (Arrow and Debreu 1954; Debreu 1959; Arrow and Hahn 1971).

The consumers include both producers and workers. We attribute to each consumer a "well-behaved" (12.5.1) neoclassical Cobb–Douglas utility function.

In his example, Sraffa has three commodities, but for one of them, iron, there is no social surplus to be distributed and consumed. The commodities to be consumed in the whole society are $c_{wheat} = 70 \quad qr.\ wheat$ and $c_{coal} = 165 \quad t.\ coal$.

Having only two consumption goods, we can characterize the decision of each consumer, as in microeconomics textbooks, in a two-dimensional space. For the sake of the exposition, we can characterize wheat as being good 1, x_1, and coal as good 2, x_2. To each agent ω we can assign endowments $\bar{x}_{1\omega}, \bar{x}_{2\omega}$. As we have shown above (Sect. 12.3.5.2), each consumer has as income a share of the value of the surplus $y_\omega^{Incomes}$ (which is an element of the vector).

We can attribute as endowments the fraction of total consumption associated to agent ω equivalent to their income. That is, $\bar{x}_{1\omega} = C_{wheat} \cdot y_\omega^{Incomes}$ and $\bar{x}_{2\omega} = C_{wheat} \cdot y_\omega^{Incomes}$.[25]

We have that $\sum_{\omega=1}^{n_\omega} \bar{x}_{1\omega} = c_{wheat} = 70$ and $\sum_{\omega=1}^{n_\omega} \bar{x}_{2\omega} = c_{coal} = 165$. As we have seen above, the individual personal incomes are related with the functional distribution of income, which is uniquely associated with a price vector and a wage rate and

[24] Provided, of course, that the associated profit rates are $\mathbf{r}^j$ and the associated wage rate is w^j. Here supply and demand meet perfectly, but the system is still indeterminate.

[25] This is done for the sake of the similarity with the standard neoclassical exchange problem where each agent is equipped with an endowment. An alternative is to assume that agents' incomes are paid with token money. This, when in bookkeeping equilibrium (as in condition b) above), does not change the substance of the results. The position of the budget constraint line depends on the income and, when prices are fixed, goes through the couple $(\bar{x}_{1\omega}, \bar{x}_{2\omega})$. The value of the endowments and that of the token money would have to be exactly the same.

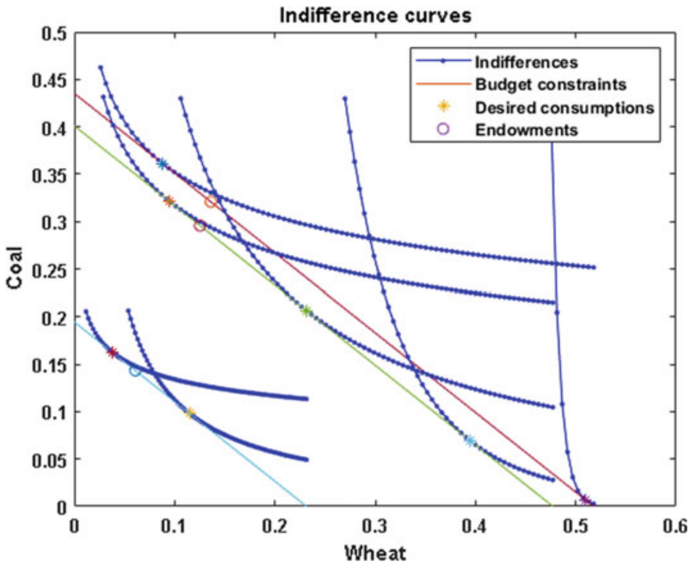


Fig. 12.1 Examples of indifference curves. These indifference and budget constraint curves are for a sample of the 300 consumers. This is the case where there is one producer in the iron industry, three in the coal industry, and fifty-six in the wheat industry. The remaining 240 are workers. This is the case with Fig. 12.2

the particular density of firms in the industries and the number of workers. Clearly, the personal incomes change when prices change.

Continuing with the construction of the textbook example, the prices $p_1 = p_{wheat}$ and $p_2 = p_{coal}$ are associated with the functional distribution of income (see the above one to one mapping, Eq. 12.4.5).

$$\max_{c_{1\omega},\, c_{2\omega}} U^{\omega} = \left(c_{1\omega}\right)^{\alpha_\omega} \left(c_{2\omega}\right)^{\beta_\omega}$$
$$s.t. \quad c_{1\omega} p_1 + c_{2\omega} p_2 = \bar{x}_{1\omega} p_1 +, \bar{x}_{2\omega} p_2 \qquad (= {}^{q}y^{j}_{\omega}) \tag{12.7.6}$$

with $\omega = 1, 2, \ldots, n_\omega = 300$ and $\alpha_\omega \in]0, 1[$ and $\beta_\omega = 1 - \alpha_\omega$.[26]

Figure 12.1 shows the budget constraints and indifference curves for some consumers.

[26] These are standard neoclassical assumptions when describing a utility function. The values α_ω were generated with the Matlab random number generator. That is, **rng(9)**, $\alpha_\omega = rand(300, 1)$, $\beta_\omega = 1 - \alpha_\omega$.

12.7.6.2 Aggregate *(excess)* Demand and Supply

The consumers' demand and supply are derived from the solution of Eq. 12.7.6.

The couple $(c_{1\omega}, c_{2\omega})$ is the desired consumption of agent ω. The difference $(c_{1\omega} - x_{1\omega})$ is the demand of agent ω for good 1 when $(c_{1\omega} > x_{1\omega})$ or is the supply of agent ω of good 1 when $(c_{1\omega} < x_{1\omega})$. The same for good 2.

The *aggregate excess demand* or *excess supply* of consumption good 1, $\overline{ads_1}$, is given by

$$\overline{ads_1} = \textit{Aggregate excess demand or supply of good 1 (wheat)} = \sum_{\omega=1}^{n_\omega} (c_{1\omega} - x_{1\omega}) \tag{12.7.7}$$

It is an *excess demand* function if greater than 0, and an *excess supply* function when less than 0.

Similarly, the aggregate excess demand or excess supply of consumption goods 2, $\overline{ads_1}$, is given by

$$\overline{ads_2} = \textit{Aggregate excess demand or supply of good 2 (coal)} = \sum_{\omega=1}^{n_\omega} (c_{2\omega} - x_{2\omega}) \tag{12.7.8}$$

The aggregate excess demand and supply, $\overline{ads}_{1,2}$, is the aggregate demand, $\overline{ad}_{1,2}$, when greater than 0, and the aggregate supply, $\overline{as}_{1,2}$ when less than 0.

12.8 Results of the Numerical Simulations

12.8.1 Virtual (Neoclassical) Economies

At this point we have a massive amount of data describing different virtual (neoclassical) economies.

An economy z is described by the following set of data:

$$\mathcal{ECONOMY}(z) = \left\{ \mathbf{d}^q, \mathbf{p}^q, \mathbf{r}^q, w^q, {}^q\mathbf{y}^q, , \mathbf{FirmsDensity}^z, {}^z\overline{ad}^q_{1,2}, {}^z\overline{as}^z_{1,2} \right\}_{q=1}^{176219} \tag{12.8.1}$$

The values $\left\{\mathbf{d}^q, \mathbf{p}^q, \mathbf{r}^q, w^q, {}^q\mathbf{y}^q\right\}_{q=1}^{176219}$ are computed just from knowing $(\mathbf{A}, \boldsymbol{\ell}, \mathbf{b})$ and the distribution of firms among industries, $\mathbf{FirmsDensity}^z$. The computation of this data does not require the description of the consumers as neoclassical utility maximizers.[27] The neoclassical component is given by the aggregate demand and

[27] Here, mainly for reasons of space and time, we abstain from discussing the conclusions one derives from the analysis of this data, $\left\{\mathbf{d}^q, \mathbf{p}^q, \mathbf{r}^q, w^q\right\}_{q=1}^{176219}$. The interested reader may inspect the figures in Zambelli (2018b) and Venkatachalam and Zambelli (2021a). The results presented in

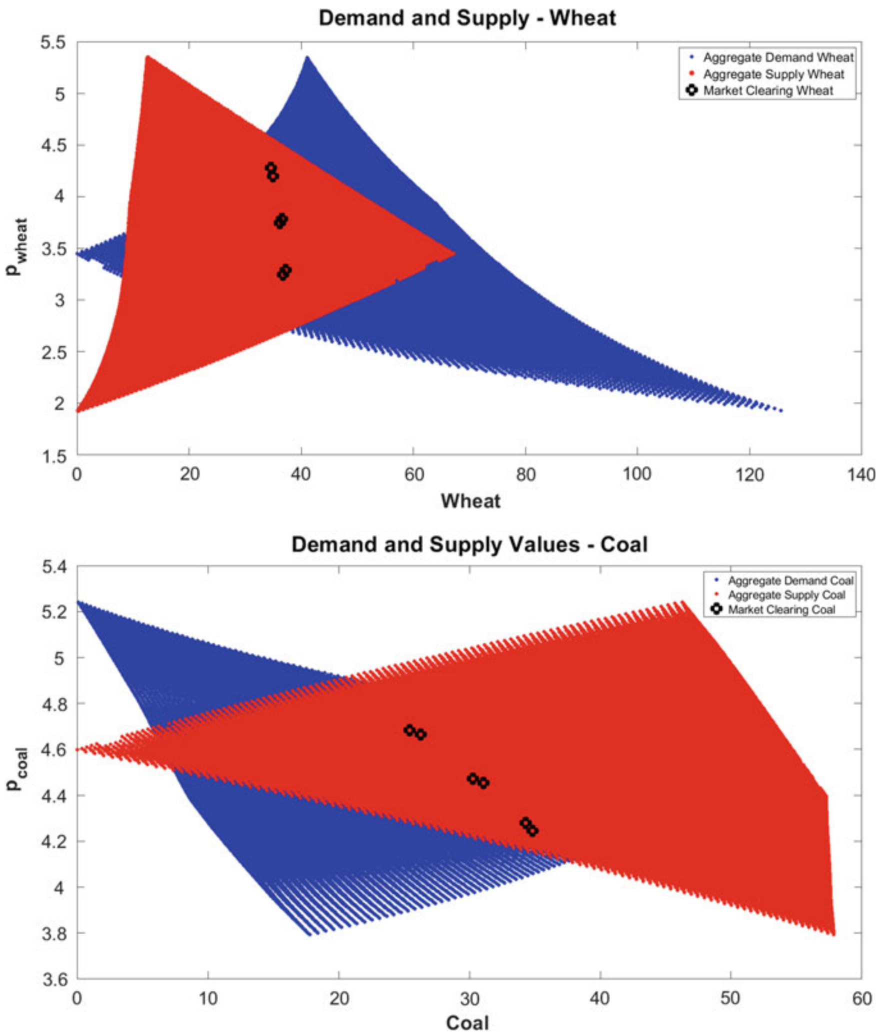


Fig. 12.2 *Aggregate demand and aggregate supply relative to* $\mathcal{ECONOMY}(z = 3)$. The upper figure is the demand and supply space $\left\{{}^z\overline{ad_1}^j, {}^z\overline{as_1}^j\right\}_{q=1}^{176219}$ for *wheat* and the lowest figure is the demand and supply space $\left\{{}^z\overline{ad_2}^j, {}^z\overline{as_2}^j\right\}_{q=1}^{176219}$ for coal for economy number $z = 3$. The $\mathbf{FirmsDensity}^{q=3} = [1, 3, 56]^T$ (one firm in the iron industry, three in the coal industry and fifty-six in the wheat industry). Nota bene. The above defines the domains of both the aggregate demand and aggregate supply. This does not mean at all that the overlapping points are equilibrium points. Actually, there are only six computed market equilibrium points (the six black dots). See also Fig. 12.4. Prices are multiplied by 1000

supply functions $\left\{ {}^{q}\overline{ad_1}^{q}, {}^{q}\,\overline{as_2}^{q} \right\}_{q=1}^{176219}$ which are the results of the attempts of the consumers to maximize their utility.

12.8.2 In Search of the Neoclassical Market Clearing Equilibria—Condition D)

There are 1711 different z and hence there are as many $\mathcal{ECONOMY}(z)$.

For each $\mathcal{ECONOMY}(z)$ we define the market clearing equilibrium for the consumption good wheat as being all the points j where:

$$\left| \; {}^{q}\overline{ad}^{q}_{wheat} \;\; - \;\; {}^{q}\overline{as}^{j}_{wheat} \; \right| < \epsilon_1 = \frac{hc_{wheat}}{n_\omega} = \frac{0.01 \times 70}{300} = 0.0023 \qquad (12.8.2)$$

The same for coal.

$$\left| \; {}^{q}\overline{ad}^{q}_{coal} \;\; - \;\; {}^{q}\overline{as}^{j}_{coal} \; \right| < \epsilon_2 = \frac{hc_{coal}}{n_\omega} = \frac{0.01 \times 165}{300} = 0.0055 \qquad (12.8.3)$$

There is an equilibrium in the economy when the above conditions hold for both markets.

The possibility of neoclassical market clearing equilibrium occurs in less than 40% of the total number of economies. When the threshold values ϵ_1 and ϵ_2 as in eqs. 12.8.2 and 12.8.3 there are only 480 (out of a total of 1711) $\mathcal{ECONOMY}(z)$ (circa 30%) where there is a market clearing equilibrium, i.e., where supply equals demand. This increases when the threshold values increase (see Fig. 12.3). With thresholds ten times greater, the number of economies where supply meets demand increases to 640 (circa 37%).

Multiple neoclassical market clearing equilibria. Of these 480 virtual economies, only 107 have only one equilibrium point, 95 have two, 81 have three, 60 have four, 43 have five, 27 have six, and on up to the greatest number of equilibria, which is thirty. The total number of equilibrium points for the whole set of economies is 1782 (out of a total number of market clearing points which is above 301 million, $301, 510, 709$: $176, 219$ functional income distributions $\{\mathbf{d}\}_{q=1}^{176219}$ times 1711 producers' distributions $\{\mathbf{FirmsDensity}^{q}\}_{q=1}^{1711}$).

Robustness of the computation of neoclassical market clearing results. The total number of economies where supply meets demand and the number of economies with multiple equilibria is a function of the approximating threshold. With a smaller threshold, the number of economies where there is market clearing diminishes, whereas

this chapter may be seen as an important addendum and/or integration of what it is presented there. However, the focus here is on the neoclassical and long-period properties of general equilibrium.

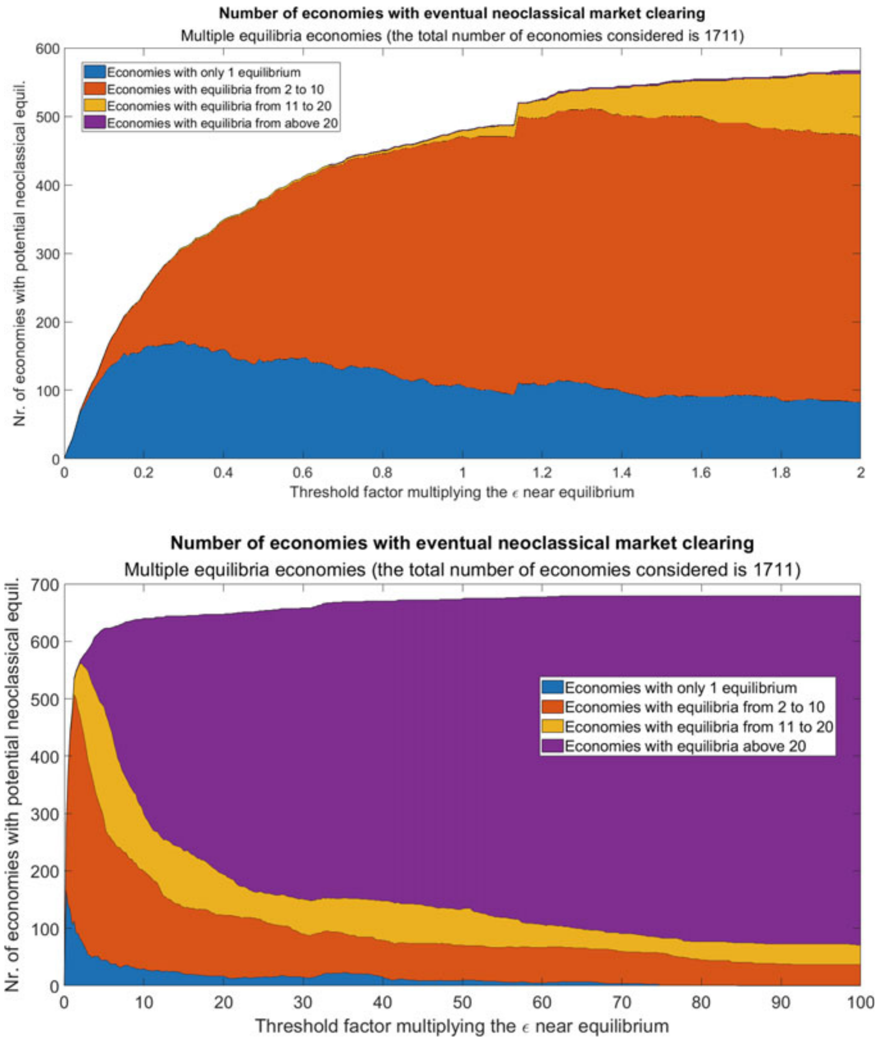


Fig. 12.3 Multiple neoclassical equilibria as the threshold values ϵ_1 and ϵ_2 change The computed number of the $\mathcal{ECONOMY}(z)$ where there is market clearing depends on the choice of the thresholds ϵ_1 and ϵ_2 (Eqs. 12.8.2 and 12.8.3). Here there is a scaling factor multiplying the values of the thresholds from 0 to 10 (with a step-size equal to 0.01). The values associated to 1 are those used in the text (Eqs. 12.8.2 and 12.8.3). The upper figure is an enlargement for the interval between 0 and of the lower figure

with a higher threshold, it increases—but up to a point. Figure 12.3 presents information concerning the number of economies where there is market clearing and also the number of equilibrium points per economy.
There is a multiplying or scaling factor of both thresholds going from 0 to 10 times the threshold values ϵ_1, ϵ_2 as in Eqs. 12.8.2 and 12.8.3. The figure shows that the number of economies which have at least one equilibrium does not grow linearly with the threshold. As is to be expected, the number of economies with multiple equilibria increases, and the number of economies with just one equilibrium decreases.

Preliminary conclusion regarding market clearing of the goods markets. Condition (d). The major conclusion, when we analyze the results of our digital laboratory, is that the number of instances where supply and demand could meet is lower than the number of instances where supply and demand would not meet. What this means is that the consumption goods markets, as well, have several degrees of indeterminacy. It is likely that either market clearing does not exist or that there are too many of them, as in the case of multiple equilibria.
Furthermore, and most importantly, one should also keep in mind that in this section we have focused on the existence of market clearing equilibria. When an equilibrium is shown to exist this does not imply at all that it is going to be realized. Let us say that market clearing in the goods market, for an $\mathcal{ECONOMY}(z)$ takes place at point $j^* \in [1, 176219]$. The associated values are $\left(\mathbf{d}^{j^*}, \mathbf{p}^{j^*}, \mathbf{r}^{j^*}, w^{j^*}, {}^q\mathbf{y}^{j^*}, \mathbf{FirmsDensity}^q, {}^q\overline{ad}^{j^*}_{1,2}, {}^q\overline{as}^{j^*}_{1,2}\right)$.
Supply is equal to demand, ${}^q\overline{as}^{j^*}_{1,2} = {}^q\overline{ad}^{j^*}_{1,2}$, if and only if the distribution among the consumers of the produced surplus is *exactly* ${}^q\mathbf{y}^{j^*}$, which implies that the functional distribution of income, the prices, the profit rates, and the wage rate must be $\mathbf{d}^{j^*}, \mathbf{p}^{j^*}, \mathbf{r}^{j^*}, w^{j^*}$, exactly.
But Sraffa's critique is that the system is indeterminate. Which means, in the context of our computations, that there is no criterion to prioritize one particular personal distribution of income, ${}^q\mathbf{y}^{j^*}$, over all possible 176,219 alternative distributions of incomes. This means that, without the imposition of new structures, any of the possible 176,219 alternative distributions of incomes have the same likelihood to occur and therefore, the likelihood of market clearing $j*$ is $\frac{1}{173219}$, which is very low.

Hence, the addition of neoclassical consumers does not render, as was to be expected, the system determinate. We have shown that when we confine ourselves to the analysis of the market clearing of the *goods markets*, even the equilibrium of these markets may not always exist, or may exhibit, like in the market for the means of production, multiple equilibria—which is another side of indeterminacy, the reason being that when we have multiple equilibria, there is no criterion that can be used to pick one of these equilibria.

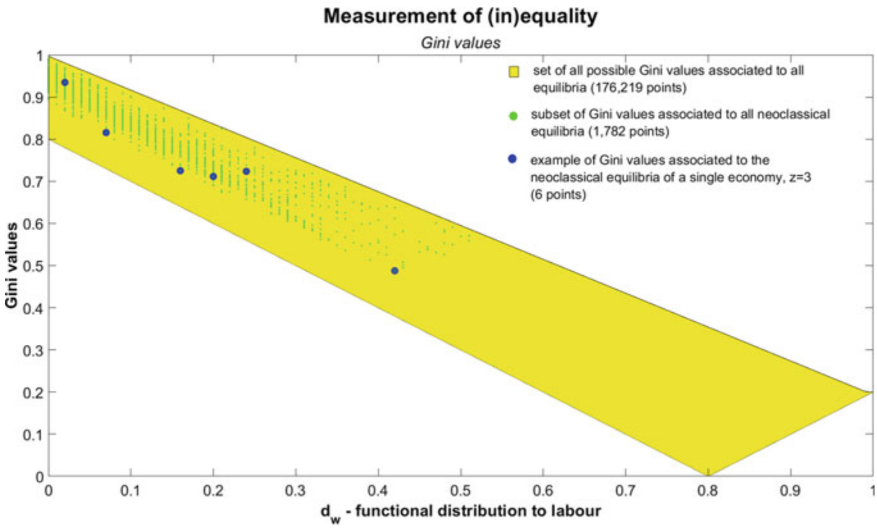


Fig. 12.4 Measurement of (in)equality: Gini values. The yellow area is the set of all Gini values associated with the distribution of income relative to the 176,219 market clearing equilibria. A Gini value equal to 1 indicates a high level of inequality (one agent gets the entire surplus). A Gini value equal to 0 indicates the highest level of equality (all agents have equal incomes). The dots or values in green are the Gini values of the personal income distribution associated with the 1782 neoclassical market clearing equilibria. The blue values are the neoclassical market clearing equilibria associated to $\mathcal{ECONOMY}(z = 3)$, see Fig. 12.2, $\textbf{FirmsDensity}^{z=3} = [1, 3, 56]^T$ (one firm in the iron industry, three in the coal industry and fifty-six in the wheat industry) The Gini values were computed with the code provided by Lengwiler (2022). Nota Bene. The abscissa reports the functional distribution of labor d_w. The worker's per capita income is $\frac{d_w}{n_\ell}$, (see Eq. 12.7.4)

12.8.3 Results Concerning the Personal Income Distributions

We use the Gini coefficient or index to measure the degrees of equality of the different distributions.

Figure 12.4 shows the association of the workers' share of the surplus with the Gini values of the associated distributions of incomes. The yellow contains all the Gini values associated with the personal distribution income vectors of all the 1711 different economic configurations that we have analyzed.[28]

Equality Clearly, given the particular composition of our population of 240 workers, 80% of the population, which are paid the same, 60 producers, 20% of the population, the lowest Gini value (highest equality among all income earners) occurs when the workers' functional distribution of income is 80%. The highest inequality occurs when the workers' functional distribution of incomes is 0 and there is only one producer receiving the whole of the surplus produced.

[28] That is, the area contains over 301 million Gini values $301510709 = 1711 \times 176219$.

The inequality of neoclassical GE market clearing The green points in Fig. 12.4 are the Gini values of the income distributions associated with the 1782 neoclassical market clearing equilibria. As is macroscopically evident, these equilibrium points indicate highly uneven distributions of incomes. It is not totally clear what might be the cause of this phenomenon. The equilibrium of supply and demand means that all the agents, given their budget constraints can buy and sell the quantities that allow the maximization of utility. In order to control that this phenomenon is not related with a particular set of utility functions, I have generated the data with different populations of agents. The result is that the lowest Gini value is the one associated with the economy reported above in Fig. 12.2 (value 0.48). The market clearing equilibria relative to the other populations are associated with higher inequality (that is, higher Gini values).[29]

As is evident from Fig. 12.2, the Gini values of the market clearing equilibrium points are far from the equality of distribution which would take place when the class of workers get 80% of the surplus.

The number of producers (60) is different from the number of workers (240). It seems that market clearing has higher chances of occurring when the distribution of income is in favor of the producers as a whole. This point deserves further investigation, but the evidence from the simulations leads to concluding that as long as the number of producers is smaller than the number of workers (which seems to be realistic), the theoretical market clearing equilibria are associated with high inequality.[30]

12.9 Conclusions

In the literature, the subtitle of Sraffa's *Production of Commodities by Means of Commodities* (PCMC), *Prelude to a critique of economic theory*, has largely been disregarded or not fully considered.

PCMC is first and foremost a critique of general equilibrium theory, but also a critique of the classical notion of the labor theory of value. Sraffa's method of study-

[29] In this context, agents are different when they are described with a different set of utility functions (i.e., a different set of α's and β's see Eq. 12.5.1). I have compared ten different populations of agents, generated with Matlab's random number generator, rng(i) with $i = 1, \ldots 10$, $\alpha_\omega = rand(300, 1)$, $\beta_\omega = 1 - \alpha_\omega$, see also footnote 26 . For each population i I have computed the same values of the major simulations scrutinized in the text. With the changes of population, we also have changes in the total number of $\mathcal{ECONOMY}(z)$ which have neoclassical market clearing equilibria, which are, as the population changes (as i changes): 88, 18, 5, 4, 15, 3, 61, 7, **480**, and 55. Among all these equilibria, the minimum Gini value associated to each distribution of incomes to each population is: 0.7, 0.73, 0.67, 0.72, 0.67, 0.82, 0.58, 0.77, **0.48**, and 0.66. The values in bold are those of the data used in the text. In passing, let me point out that the economy used is the one, among the ten different one analyzed, which is more favorable to the neoclassical market clearing equilibria because it has the highest number of neoclassical market clearing equilibria (480) and the lowest Gini value (0.48).

[30] Hence, these equilibria provide the potential for a class struggle or may lead to the "instability of capitalism" Schumpeter (1928).

ing the properties of self-replacing prices, profit rates, and wage rates, is powerful because it does apply to any notion of economic bookkeeping equilibrium (where the revenues of all the agents must be equal to their expenditures). Sraffa's method allows avoiding studying the theories describing individual motivations, and the specifics of the methods of production available to the individual firms. Therefore it applies to a variety of equilibrium positions.

An economic general equilibrium is simply a condition that repeats itself (because there is by definition no change[31]). Sraffa concluded that the self-replacing prices are not *computable* based on no more than the knowledge of the methods of production.

The relevance and breadth of the critique are valid, and are in my view amplified, when self-replacing prices are not confined to the narrow distribution set associated to a uniform rate of profits (Zambelli 2018b; Venkatachalam and Zambelli 2021a, b).

PCMC is a reinforcement of the critiques that Sraffa personally addressed during the late 20s and the 30s toward the notions of general equilibrium of his contemporaries Hayek and Hicks. But also a critique to the general equilibrium notions developed in the tradition of Walras, Pareto, Edgeworth Böhm-Bawerk, Wicksell and Marshall (partial equilibrium). That is, the neo-classical equilibrium where supply equals demand is not determined by the forces of free competition. In other words, general equilibrium is not determined by *economic forces* of *free* or *perfect competition*. Other elements have to be considered.[32]

The neo-Walrasians have merely circumvented the problem of indeterminacy by exogenously fixing the prices.

With respect to the *neo-classical* notion of general equilibrium, the conclusion is that market clearing equilibria in the market for consumption goods do not occur for most of the economies we have considered (with different distributions of firms among the industries). In the very few instances, associated with specific distributions of firms, where aggregate supply meets aggregate demand for the consumption goods, we have shown that these instances are accompanied with multiple equilibria.

Therefore the problem of indeterminacy – defined broadly as the condition where there is no solution or where there are multiple solutions – is strengthened by the fact that when agents (both as consumers and producers) are added to the Sraffian schemes, so as to have supply and demand of consumption goods, the so-called "law" that supply meets demand seems not to hold for most cases. When it holds, the

[31] The models that study growth in a balanced growth path that postulate Walrasian general equilibrium imply a change of the quantities produced and demand, but postulate also that at each point in time (or interval of time) there is a static Walrasian general equilibrium of the Arrow-Debreu type. Obviously a static system is a system without change.

[32] Many readers will find this result at odds with the literature, which asserts to have "demonstrated" the existence of general equilibrium (Arrow and Debreu 1954; Debreu 1959; Arrow and Hahn 1971; Scarf 1973). Here we have used the precision of the algorithmic method of investigation. The algorithms reflect the equations presented above. They are simple (i.e., require the writing of a few lines of code). The robustness and generality of the results is in my view evident, but anyone who wishes to do challenge them by constructing counter-examples. It is my conviction that counter-examples can not be produced.

possibility of multiple equilibria excludes the possibility of computing the solution with endogenous mechanical economic forces.

It has to be added that these neoclassical market equilibria are also associated to extreme inequality of income distribution. The essence of Sraffa's critique presented in PCMC was already present in Sraffa (1926). What Sraffa demonstrated in 1960 is that *the essential causes determining the price of particular commodities may not be simplified and grouped together so as to be represented by a pair of intersecting curves of collective demand and supply*.[33]

Sraffa's critique, that the aggregate demand and the aggregate supply almost never meet, is devastating for any 'old' (Walras, Böhm-Bawerk, Baldone, Jevons, Marshall, and so on) or 'modern' (Hayek, Hicks, Arrow, Debreu, Hahn, and so on) marginalist general equilibrium system.

The postulate that the economic system is in a perpetual condition of general equilibrium should be removed from economic theorizing.

Appendix

12.10 Sraffa's View on the Incapacity of the Economic System to Endogenously Determine Distribution

As we have already pointed out in Venkatachalam and Zambelli (2021b, pp. 467–468), in his correspondence with Garegnani, Sraffa objected to Garegnani's interpretation of the hint given by Sraffa's PCMC that the "The rate of profits […] is accordingly susceptible of being determined from outside the system of production, in particular by the level of the money rates of interest." (PCMC, p. 33)

Sraffa was clearly also concerned about the possibility of changing the distribution. His correspondence with Garegnani is very illuminating.

Sraffa's letter to Garegnani and the personal handwritten notes on Garegnani's letter show a total disagreement with respect to matters that Garegnani had written in his letter. For example, Sraffa writes, commenting on some of Garegnani's statements (Sraffa Papers, D3/12/111/142.1.recto):

> quello che voi sempre ignorate (da Marshall in poi) e' che i metodi di produzione cambiano tutto il tempo anche senza sia pur piccole variazioni della quantita' prodotta !
>
> [My translation: what you all ignore (from Marshall onwards) is that production methods change all the time even without variations in the quantities produced] (Sraffa Papers, D3/12/111/142.1.verso)

Clearly, if *production methods change all the time*, the method related with long-period analysis is unimportant if based on unchanged methods.

[33] See above the quotation at page 208 taken from Sraffa (1926, p. 535).

In the same page as a comment to Garegnani, Sraffa writes (ironically):

> evidentemente 'l'edificio' che hai in mente e' quello di Marshall, il ponte che si regge sui due pilastri della domanda e dell'offerta
>
> [My translation: evidently 'the building' you have in mind is that of Marshall, the bridge that rests on the two pillars of supply and demand] (Sraffa Papers, D3/12/111/142.1.recto)

Here we have some additional evidence for the fact that Sraffa's work was a critique of economic theory and in particular of the idea that supply and demand intersect. asked some clarifications and provided a possible mechanical interpretation that the interest rate could be determined in the money market [Sraffa Papers, D3/12/144/145].

Clearly, if one follows Garegnani's suggested interpretation, one could erroneously infer that the money interest rate(s) could by itself determine the profit rate(s). And given that the interest rate(s) may in turn be determined by the supply and demand of funds, the self-replacing prices of the economic system would also be determinate by the *mechanical forces of the markets*.

Sraffa responded firmly to Garegnani's interpretation:

> …I have no intention of putting forward yet another mechanical theory which, in one form or another, reinforces the idea **that distribution is determined by natural, or technical, or perhaps even accidental circumstances**, but **such as to render futile whatever action**, from one side or the other, **aimed to modify the distribution**. (Sraffa Papers, D3/12/111/149, my translation, in bold emphasis added)

12.11 Existence of Equilibrium

The question of whether in a capitalistic system of production *market clearing* equilibrium prices exist and can be determined by agents' interactions was originally studied by *marginalist* economists as being isomorphic to the problem of finding the solution of a system of equations, which meant checking whether the number of equations was the same as the number of unknowns (see Walras (1874, pp. 43–44), Pareto (1906, Chap. 3, pp. 297–99) and many others, including some aspects of the work of Edgeworth, Böhm-Bawerk, Wicksell, Marshall, Jevons and many others).

After World War II (preceded by Hicks (1939) and Wald (1951)), Arrow and Debreu (1954); Debreu (1959); Arrow and Hahn (1971) replaced the method of counting equations and unknowns by studying the same problem by means of the mathematics of convex, compact and closed sets (whose elements belong to the real hyperspace $\mathbb{R}^n$ and the functions defined on these sets are continuous and mostly differentiable). The existence of the equilibrium is then presumed to be proved by applying the Brouwer fixed-point theorem. See Debreu (1959, Chap. 1) or Arrow and Hahn (1971, Chap. 2).

Velupillai (2006, Theorem 2, p. 363) has demonstrated that the existence of the Walrasian general equilibrium prices as defined in Arrow and Debreu (1954) "is

undecidable, i.e., cannot be determined algorithmically." Therefore, the existence of equilibrium prices, cannot, in general be proven.[34]
But also the method of counting the equations and the unknown does not prove existence simply because the system of equations may not be solvable. What this method provides is only a criterion for the identification of a necessary condition for the solution to be found, surely not a sufficient one. On the specific issue of the computability of equilibria, see the collected articles in (Velupillai et al. 2011, Part V. *Computable General Equilibrium Theory*. Part VII. *Computable Microeconimics*)

12.12 Sources of Personal Income

Matrix **H** (Eq. 12.3.18, page 216) contains all the information concerning the source of each agent personal distribution of income. Clearly the income of an agent ω is given by the sum of all that agent's income (work as well as dividends from profits):

$$\textit{Income of individual}\omega = \mathbf{e}^T_{(n+1)\times 1}\mathbf{h}^{\omega}\mathbf{s}^T\mathbf{p} = (\sum_{i=1}^{n+1} h_i^{\omega})\mathbf{s}^T\mathbf{p}$$

which is, obviously, a scalar.

The sum by columns of matrix **H** is the share of the society's income going to the individuals:

$$\bar{\mathbf{h}}^T = \textit{Distribution of income among individuals} = \mathbf{e}^T_{(n+1)\times 1}\mathbf{H}\mathbf{s}^T\mathbf{p}$$

The vector $\bar{\mathbf{h}}$ is the share of the income to the individual, personal income. It is composed of n_ω elements. The elements are positive, when self-replacing is taking place, and their sum is equal to 1.

The sum of all the elements of **H** is 1.

The sum by rows of **H** is the *functional distribution of income*:

$$\mathbf{d} = \mathbf{H}\mathbf{e}_{n_\omega \times 1} \qquad (12.12.1)$$

The sum by columns of matrix **H** is the share of total income going to individual agents, **y**, and (when *numéraire* is the surplus, $\mathbf{s}^T\mathbf{p} = 1$) it is also the *personal income* per agent, $\mathbf{y}^{Incomes}$:

---

[34] Velupillai has discussed the fundamental question of the computability of general equilibrium in several contributions. (Velupillai, 2005, pp. 860–63) discusses in depth the limits of finding the solutions to a system of equations even when the number of equations and unknown are the same and, most importantly, the issue of whether the Arrow–Debreu–Hahn mathematics is constructive, i.e., allows the actual computation of the equilibrium points. His clear conclusion is that not even the claim that general equilibrium may be algorithmically found Scarf (1973) is correct. Therefore, the claim that Scarf's algorithm may be used to 'compute' that Computable General Equilibrium model is empty. See also (Velupillai 2010, Part II, *General Equilibrium Theory* and Part III *Methodology*).

$$
\begin{aligned}
\mathbf{y} &= \mathbf{H}^T \mathbf{e}_{(n+1)\times 1} \\
\mathbf{y}^{Incomes} &= \mathbf{H}^T \begin{bmatrix} \mathbf{p} \\ w \end{bmatrix} \\
\mathbf{y}^{Incomes} &= \mathbf{y}
\end{aligned}
\tag{12.12.2}
$$

This equation is important because it encapsulates, when self-replacing is taking place, the relation, *in physical terms as well as in value terms*, existing between the functional distribution of income and the personal distribution of income.

As we have done above when considering individual consumption (Sect. 12.3.3.1) we may divide the total population n_ω of income earners into two groups: n_π producers and n_ℓ workers.

Consequently, we can partition $\mathbf{H}$ into a matrix containing all the information regarding the sources of income of the producers, ${}^{n_\pi}\mathbf{H}_{(n+1)\times n_\pi}$, and another describing all the sources of income of the workers, ${}^{n_\ell}\mathbf{H}_{(n+1)\times n_\ell}$.

That is:

$$
\mathbf{H} = \left[{}^{n_\pi}\mathbf{H} \;:\; {}^{n_\ell}\mathbf{H} \right] \tag{12.12.3}
$$

Consequently,

$$
\mathbf{h} = \mathbf{h}^\pi + \mathbf{h}^\ell = {}^{n_\pi}\mathbf{H}\mathbf{e}_{n_\pi \times 1} + {}^{n_\ell}\mathbf{H}\mathbf{e}_{n_\ell \times 1} \tag{12.12.4}
$$

or, which is the same,

$$
\mathbf{d} = \mathbf{d}^\pi + \mathbf{d}^\ell = {}^{n_\pi}\mathbf{H}\mathbf{e}_{n_\pi \times 1} + {}^{n_\ell}\mathbf{H}\mathbf{e}_{n_\ell \times 1} \qquad (12.C.4a) \tag{12.12.5}
$$

where $\mathbf{d}^\pi + \mathbf{d}^\ell$ is a vector of dimension $(n+1)\times 1$ of the share of the surplus $\mathbf{s}$ going to the producers and $\mathbf{d}^\pi + \mathbf{d}^\ell$ is the share of the surplus $\mathbf{s}$ paid to the workers.

If we assume that the only source of income of the producers is dividends from the firms they own, and the only income of the workers is from the labor they sell, then, by combining Eq. 12.12.5 with Eq. 12.3.17 we can see that:

$$
\mathbf{d}_{(n+1)\times 1} = \mathbf{d}^\pi + \mathbf{d}^\ell = \begin{bmatrix} \mathbf{RAp} \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{n\times 1} \\ \mathbf{e}^T \boldsymbol{\ell} w \end{bmatrix} \tag{12.12.6}
$$

The personal income may also be organized taking in consideration the two groups:

$$
\mathbf{y}^{Incomes} = \begin{bmatrix} \mathbf{y}^{Incomes\ Producers} \\ \mathbf{y}^{Incomes\ Workers} \end{bmatrix} = \begin{bmatrix} {}^{n_\pi}\mathbf{H}^T \\ \hline {}^{n_\ell}\mathbf{H}^T \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ w \end{bmatrix} \tag{12.12.7}
$$

12.13 Appendix. Distribution, Prices and Wage Rate once the Profit Rates are Given

In the previous section we have explained how to generate the space or set of all possible self-replacing solutions. If we were given an arbitrarily picked vector of profit rates, we could check whether it belongs to the set of solutions by searching for an equal vector from the space of solutions.

But there is an alternative, in general and for any given *numéraire*, $\boldsymbol{\eta}$, when the vector of profit rates, $\mathbf{r}$ is given there is a vector of prices $\mathbf{p}$ and a wage rate w that could be associated to it by putting relations Eqs. 12.3.15 and 12.3.16 together:

$$\left[\begin{array}{c|c} \mathbf{B}-(\mathbf{I}+\mathbf{R})\mathbf{A} & -\mathbf{L} \\ \hline \boldsymbol{\eta}^T & 0 \end{array}\right]\left[\begin{array}{c} \mathbf{p} \\ \hline w \end{array}\right]=\left[\begin{array}{c} \mathbf{0}_{n\times 1} \\ \hline 1 \end{array}\right] \tag{12.13.8}$$

As we discussed in the previous section for the case of arbitrary prices, it is not the case that arbitrary profit rates would be associated with self-replacing profit rates.

A price vector necessary to allow the system to reproduce the previous year production would be

$$\left[\begin{array}{c} \mathbf{p_r} \\ \hline w_\mathbf{r} \end{array}\right]=\left[\begin{array}{c|c} \mathbf{B}-(\mathbf{I}+\mathbf{R})\mathbf{A} & -\mathbf{L} \\ \hline \boldsymbol{\eta}^T & 0 \end{array}\right]^{-1}\left[\begin{array}{c} \mathbf{0}_{n\times 1} \\ \hline 1 \end{array}\right] \tag{12.13.9}$$

The price vector and wage rate (when we expand the inverted matrix) would be respectively

$$\mathbf{p_r}=\left[\mathbf{B}-(\mathbf{I}+\mathbf{R})\mathbf{A}\right]^{-\mathbf{1}}\mathbf{L}w_\mathbf{r} \tag{12.13.10}$$

and

$$w_\mathbf{r}=\frac{1}{\boldsymbol{\eta}^\mathbf{T}(\mathbf{B}-(\mathbf{I}+\mathbf{R})\mathbf{A}))^{-\mathbf{1}}\mathbf{L}} \tag{12.13.11}$$

The distribution of the surplus associated to these prices and wage rate would be

$$\mathbf{d_r}\mathbf{s}^T\mathbf{p}=\left[\begin{array}{c} (\mathbf{B}-\mathbf{A})\mathbf{p_r}-\mathbf{L}w_\mathbf{r} \\ \mathbf{e}^T\mathbf{L}w_\mathbf{r} \end{array}\right] \tag{12.13.12}$$

and $\left(d_z \geq 0 \quad z \in [1, n+1]\right)$ and $\left(\sum_{q=1}^{n+1} d_z = 1\right)$. Recall that $\mathbf{S}^T\mathbf{p}$ is the Net National Product measured at values determined by the *numéraire*, $\boldsymbol{\eta}$. The value of the surplus going to the producers as a whole is

$$\textit{Value of s to Producers} = \sum_{i=1}^{n}\left(d_\mathbf{r}(i)\mathbf{s}^T\mathbf{p_r}\right)=\left(\mathbf{s}^T\mathbf{p_r}\right)\sum_{i=1}^{n}d_\mathbf{r}(i)=\left(\mathbf{s}^T\mathbf{p_r}\right)d_\pi \tag{12.13.13}$$

where $d_\pi = \sum_{i=1}^{n} d_\mathbf{r}(i)$.

Consequently the value of the surplus going to the workers as a whole is

$$Value\ of\ s\ to\ Workers = d_w \mathbf{s}^T \mathbf{p_r} \tag{12.13.14}$$

Clearly

$$\mathbf{s}^T \mathbf{p_r} = d_\pi \mathbf{s}^T \mathbf{p_r} + d_w \mathbf{s}^T \mathbf{p_r} \tag{12.13.15}$$

The above relations, shares and prices would hold for any given *numéraire*. If we pick the *numéraire* to be the surplus $\boldsymbol{\eta} = \mathbf{s} = (\mathbf{B} - \mathbf{A})^T \mathbf{e}$, we have that $\mathbf{s}^T \mathbf{p} = 1$ and substituting these values into Eqs. 12.13.9, 12.13.10, and 12.13.11, we obtain a simplified Eq. 12.13.12, where the distribution of the surplus between industries and labor and the absolute values of the surplus coincide numerically:

$$\mathbf{d_r} = \begin{bmatrix} (\mathbf{B} - \mathbf{A})\mathbf{p_r} - \mathbf{L}w_\mathbf{r} \\ \mathbf{e}^T \mathbf{L} w_\mathbf{r} \end{bmatrix} = \mathbf{d_r} \mathbf{s}^T \mathbf{p} \tag{12.13.16}$$

Clearly $d_\pi = 1 - d_w = \mathbf{e}^T \left[(\mathbf{B} - \mathbf{A})\mathbf{p_r} - \mathbf{L}w_\mathbf{r} \right]$, independently of the *numéraire*.

References

Aberth O (2007) Introduction to precise numerical methods. Academic Press, Elsevier, Amsterdam

Arrow K, Hahn F (1971) General competitive analysis. Holden-Day

Arrow KJ, Debreu G (1954) Existence of an equilibrium for a competitive economy. Econometrica 22(3):265–290

Bharadwaj K (1963) Value through exogenous distribution. In: The Economic Weekly, pp 1450–1454

Bracci G (2022) Piero Sraffa and monetary theory: a reappraisal. Dissertation Paper-Chapter 2-University of Trento, pp 1–40

Debreu G (1959) Theory of value. Wiley

Debreu G (1987) Existence of general equillibrium. In: Eatwell J, Milgate M, Newman P (eds) The New Palgrave: a dictionary of economics, vol 2. MacMillan and Stockton, pp 216–218

Donzelli F (2006) Walras and Pareto on the meaning of the solution concept in general equilibrium theory. Rivista Internazionale di Scienze Economiche e Sociali 53(4):491–530

Hayek F (1931) Prices and production. Routledge, London

Hicks J (1939) Value and capital, 2nd edn. Oxford University Press, London. (1946)

Lengwiler Y (2022) Gini coefficient and the Lorentz curve. MATLAB Central File Exchange. Accessed 21 Jan 2022

Mandler M (1999) Sraffian indeterminacy in General Equilibrium. Rev Econ Stud 66(3):693–711

Marshall A (1890). Principles of economics, 8th edn. Macmillan, London. (Printed in 1920. Sraffa owned and annotated a 1922 reprint, 1920)

Pareto V (1906). In: Montesano A, Zanni A, Bruni L, Chipman JS, McLure M (eds) Manual of political economy: a variorum translation and critical edition. Oxford University Press. (2014)

Scarf H (1973) The computation of economic equilibria. Yale University Press

Scarf H (1987) Computation of general equilibrium. In: Eatwell J, Milgate M, Newman P (eds) The New Palgrave: a dictionary of economics, vol 1. Macmillan and Stockton, pp 556–561

Schumpeter J (1928) The instability of capitalism. Econ J 38(151):361–386

Sraffa P (1925) Sulle relazioni tra costo e quantità prodotta. Annali di Economia II:277–328

Sraffa P (1926) Law of returns under competitive conditions. Econ J 36(144):535–550

Sraffa P (1932) Dr. Hayek on money and capital. Econ J 42(165):42–53

Sraffa P (1960) Production of commodities by means of commodities. Cambridge University Press, Cambridge

Velupillai KV (1989) The existence of the standard system: Sraffa's constructive proof. Polit Econ 5(1):3–12. (Studies in the Surplus Approach)

Velupillai KV (2005) The unreasonable ineffectiveness of mathematics in economics. Camb J Econ 29:849–872

Velupillai KV (2006) Algorithmic foundations of computable general equilibrium theory. Appl Math Comput 179:360–369

Velupillai KV (2010) Computable foundations for economics. Routledge, London

Velupillai KV, Zambelli S, Kinsella S (2011) Computable economics, Edward Elgar, Cheltenham

Venkatachalam R, Zambelli S (2021a). Self-replacing prices with credit and debt. Struct Chang Econ Dyn 60(1):451–466

Venkatachalam R, Zambelli S (2021) Sraffa, money and distribution. In: Sinha A (ed) A reflection on Sraffa's revolution in economic theory, chapter 15. Palgrave-Macmillan, London, pp 455–490

Wald A (1951) On some systems of equations of mathematical economics. Econometrica 19(4):368–403

Walras L (1874) Elements of pure economics. Allen and Unwin, London. (A translation of the Edition Définitive (1926) of the (Elements d'economie politique pure), annotated and collated with the previous editions. Editor-translator, Jaffe W, 1954)

Zambelli S (2018) The aggregate production function is NOT neoclassical. Camb J Econ 42:383–426

Zambelli S (2018) Production of commodities by means of commodities and non-uniform rates of profits. Metroeconomica 69:791–818

Chapter 13
Visualizing the Roles of Frequent Terms in LTEs Following Two Economic Crisis Trigger Articles

Brian J. Kokensparger

Abstract The Letters to the Editor (LTEs) written over a twelve-week period after articles reporting the likelihood of sustained impact of two oncoming economic crises (the Oil Crisis of October 1973, and the COVID-19 Crisis beginning in March of 2020) were analyzed for the frequency of terms in both total usage as well as usage in unique LTEs. These terms were scored by multiplying total occurrences with occurrences in unique LTEs and then ranked by score for each of the 12 weeks of the study period. The findings produced sports-team-like ranking-induced shapes along the weekly rankings that can be visually compared with other terms in the same corpus, as well as the same and different terms in other corpora. This study revealed that some of the same terms seemed to have similar curve shapes in both corpora. The study also revealed that some terms, though different between the two corpora, appeared to have the same curve shapes—and perhaps may fulfill similar roles in their respective LTE discourses.

Keywords Economic crisis · Ranked data · Pandemic · Oil crisis · LTE discourse

13.1 Introduction

On October 18th, 1973, Omahans opened the morning edition of the *Omaha World-Herald* to read "Arab oil ministers in Kuwait Wednesday announced they will cut production at least 5 per cent a month until Israel gives up all the territory it seized in the 1967 war and restores the rights of the Palestinian refugees." (OWH page 2).

On March 3rd, 2020, Omahans opened up that day's edition of the paper to read, "Ricketts Likens Virus Preparation to Stocking Up for a Snowstorm" (OWH, page 1), and "Virus: Cancellation of large-scale events possible" (OWH page 3).

B. J. Kokensparger (✉)
Computer Science, Design and Journalism Department, Creighton University, 2500 California Plaza, Omaha, NE 68178, USA
e-mail: bkoken@creighton.edu; BrianKokensparger@creighton.edu

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_13

In both cases, these news stories triggered the beginning of what would eventually prompt a host of Letters to the Editor (LTEs) to the *Omaha World-Herald*, discussing a myriad of aspects of the crises.

LTEs are popular unsolicited short-format opinion pieces submitted to newspapers, magazines, and journals for publication from among the publication's general readership. These publications' opinion editors select a small subset of the letters that are submitted, based on their own internal policies and protocols. As a data source, LTEs are akin to "Vox Populi" (roughly defined as "voice of the people"), so they often include subjects and sentiments that are not nuanced by editorial control, and thus allow readers to get a feel for what "the people" are discussing in diners and street corners, barber shops and hair salons, and around the water cooler at workplaces. In terms of narratives surrounding an event, LTEs provide a cross section of sentiments that provide more depth and variation to the mainstream rhetoric (though they sometimes parrot parts of the rhetoric, often out of its original context, if it suits the writer's point).

In regard to economic crises, policymakers and business executives need public support to enact policies and other solutions in dealing with the crises, and LTEs are good indicators of what level of support they have "in hand" for their policies and solutions. The LTEs are also instructive in allowing public relations professionals to devise strategies for implementation that take public opinion into account, providing a higher chance for a successful campaign.

Economic crises often are first introduced to the general public via publication of "trigger stories" through media outlets in print, audio or video formats. A trigger story is a media story that is identified (well into the crisis or after the crisis is over) as being a trigger event that sets the crisis into motion; it is often marked as the attributed beginning of the crisis.

Two crises were chosen for this study: The Oil Crisis of 1973, and the COVID-19 Pandemic of 2020 (referred to from hereon as the COVID Pandemic). They were chosen for these reasons:

- They both affected citizens of the United States, Nebraska, and Omaha.
- They are both modern crises, and thus (except for the presence of the Internet and other digital technologies on the latter crisis) the industrial and health care contexts are similar.
- They are of very different subjects (energy and health), so one would not expect very many terms to be common between them. For example, the subject of energy might provoke discussions of solar power, whereas the subject of health might provoke discussions of vaccinations and treatment. To expect the energy discussions to include a term like "vaccinations" is as ludicrous as expecting health discussions to include "solar panels" as a frequently-used term.

Both crises were addressed in the "Public Pulse", an LTE section on the opinion page of the *Omaha World-Herald*, Nebraska's daily newspaper since 1885 (also as the *Omaha Daily-Herald* since 1878) (*Omaha World-Herald* 1973, 2020). The opinions expressed in the LTEs are records relating to world and U.S. events about the crises as seen through the eyes of citizens of Omaha (and nearby towns) and mostly through print and broadcast news outlets for the Oil Crisis, with the addition of cable and Internet news outlets for the COVID Pandemic. For both crises, 12 weeks of responses (in the form of LTEs) by the everyday readers of the *Omaha World-Herald* were recorded and analyzed in this study.

Why do Term Frequency Analysis at all? Why is it important? The brevity of the LTEs requires writers to get right to the point with their opinions, and they must also write about matters that are the most important to them. Unlike an opinion essay, which often is written with a broad view of the problem and balance, an LTE is a narrow view of the problem and often espouses only one impassioned view. Being able to categorize the specific subject of a given LTE through Term Frequency Analysis allows us to discover what is truly important to the people. For example, in the Oil Crisis of 1973, was its effect on lowering the speed limit on highways more important to the average citizen than rationing gasoline? Term Frequency Analysis can give us insight into this question, and others. During the COVID-19 Pandemic of 2020, was its effect on forcing students out of the traditional classroom more important to the average citizen than requiring all people to wear masks in public buildings and businesses?

This study is focused on performing Term Frequency Analysis on LTEs published after trigger articles about the two crises were published in the *Omaha World-Herald*. The hope is that this study will be a first step in beginning to predict terms that will have significant occurrence patterns in LTEs that are published the first few weeks after a crisis arises.

13.2 Literature Review

Shiller brought the field of narrative economics into the mainstream with his famous paper by that name (Shiller 2017). Recent work has furthered the focus on narrative economics particularly in relation to markets (Paugam et al. 2020; Holmes 2019). In a doctoral dissertation further exploring the field, Volchik (2017) states "the analysis of narratives may precede the economic and institutional analysis of crises and anti-crisis policies" Volchik adds, "Pioneering works of R. Shiller, G. Akerlof and D. Snower spotlighted the importance of analyzing narratives and narrative influence when studying economic processes." Shiller seems to extend this idea in a later paper (Shiller 2020).

In addition to the concept of narrative economics, Koslyk (2019) presents narrative approaches to news cycles, which directly relates to this study. Soroka et al. (2014) discussed media coverage and how it reflects economic trends. Barsky & Sims (2011) introduced the idea of "news shock" tied into business cycles and explored additional dynamics of news shock in a later paper (Barsky, Basu, & Lee, 2015). Alexopoulos and Cohen (2015) applied these dynamics to market volatility and decreased market returns. Although this paper does not examine news shocks, per se, there is some evidence to suggest that these media stories provide a similar dynamic to those described by Barsky, et al., in the two papers.

Besides the shock value related to news cycles, the content of news cycles may be mined using standard textual content analysis techniques, such as word frequencies, collocation (the nearness of one word to other related words), sentiment, etc. Dyas-Correia and Alexopoulos (2014) used data mining to look for “nuggets” in text. Also Gentzkow et al. (2019) examined text as data and provided a workflow for using text to inform economic research, including a discussion of relevant statistical methods and applications.

Writing this paper as the COVID-19 pandemic is ongoing, there are a plethora of articles outlining the origins of viral infections in the United States, and articles specifically mentioning infections in Omaha, Nebraska. The *Omaha World-Herald* digital download website on Newsbank.com provides access to these articles, many significant ones which will most likely be published after this paper is in the various publication stages.

Dr. Shu-Heng Chen (2012) described “varieties” of agents in his exemplary work in agent-based computational economics (ACE), which informed many of the theoretical choices made in this paper. As LTEs affect the content of other LTEs published shortly after them, there is a real sense of social interaction (albeit in the printed word with the latency of writing, submission, acceptance, and typesetting times), and therefore some of the theoretical framework around swarming and other cellular-automata phenomena may be satisfactorily applied. The decision to accept one LTE over another also lies within the purview of the editorial staff, which makes choices based on space limitations in the LTE section of the newspaper. Although these editorial staff members would not in any way be considered zero-intelligence human agents, they often must put aside their knowledge and publish LTEs that reflect naïve and ignorant viewpoints for the sake of balance. Since the LTE section of the newspaper purports to reflect the “Public Pulse” of the people (and therein lies its popularity and affects marketability of the newspaper), one hundred LTE submissions with an ignorant viewpoint cannot be ignored and replaced by one LTE submission with supported factual reasoning and cogent argumentation. Therefore, the LTE dynamic more appropriately reflects zero-intelligence agents but strives for near zero-intelligence.

Similarly, Hsu, Yu & Chen approached an economic problem (the U.S. Silver Purchase Act and its effect on the Chinese price level of silver) with textual analysis methods that were also applied to narrative frequencies of terms in newspapers (2021). The data were drawn from the Shanghai *ShunBao* newspaper database (Hsu et al. 2021. P. 9), which is similar to the newspaper used in this study (the *Omaha World-Herald*) in that it is published for general readership and includes some national and international news. The textual analysis methodology for the Silver Purchase Act study is quite different than that employed in this study, which is reasonable because the focus of the Hsu, Yu & Chen study is on the effect of narrative keywords on the Chinese price level (with the example of their effect on the Shanghai WPI), whereas the focus of this study is on the trending of keywords either up or down within the discourse during the study periods after the trigger stories of the explored crises.

13.3 Methodology

The *Omaha World-Herald* provides a digital archive of images of all its daily editions from 1885 to the present. Every issue includes an opinion page, with a "Public Pulse" section, usually including from 3 to 12 LTEs. For LTEs to be considered for publication in the Public Pulse, they must include full contact information (though that information may be withheld from publication by request), they must be signed, and they must not plagiarize the work of others. Also, the editorial staff of the Public Pulse reserves the right to edit letters for reasons of "taste, accuracy, clarity and length." (OWH, April 15, 2020).

For each of the two triggering events, 12 weeks of daily newspaper editions of the *Omaha World-Herald* were manually inspected for general key words or phrases that tied the individual letter in some way to the triggering event. During this data collection period, the reader looked for certain key words to aid in the manual process. For the Oil Crisis, any LTE that included such terms as "energy," "fuel," "oil," any of the terms associated with "gasoline," and other energy-related terms discussing the shortage, usage, conservation, or management of energy resources were selected. For the COVID Pandemic, general key words that tied the individual letters in any way to the COVID-19 virus, or any other item related to the virus, such as quarantine restrictions, mask mandates, and local, state, and national lockdowns, were included.

13.3.1 Generating LTE Corpora

For both corpora, a snip image was generated for each selected LTE and saved in a yyyymmdd_99.png format. The LTEs in these snip images were transcribed into text documents bearing the same naming format. A proofing algorithm was employed to find and correct typographical entry errors. These text files were saved in corpus folders for each of the two focus crises.

230 LTEs were collected for the Oil Crisis during the 12-week study period, and 201 were collected for the 12-week study period during the COVID Pandemic.

The weekly total number of LTEs with total word count for both study periods are provided in Fig. 13.1.

In regard to the number of files per week, both corpora have a similar m-curve, with the Oil Crisis files being generally higher in number than the COVID Crisis files. However, in regard to the total words by week, the two curves are roughly similar—a little more like a normal curve than the number of files. In total words per week, however, the COVID Crisis had much higher values. This is most likely due to editorial decisions: in 1973, the apparent editorial choice was to publish more LTEs, but limit their word counts. In 2020, the apparent editorial choice was to include fewer LTEs, but to publish LTEs with a higher word count. Since both corpora were analyzed and ranked separately within each individual corpus (i.e., there was no

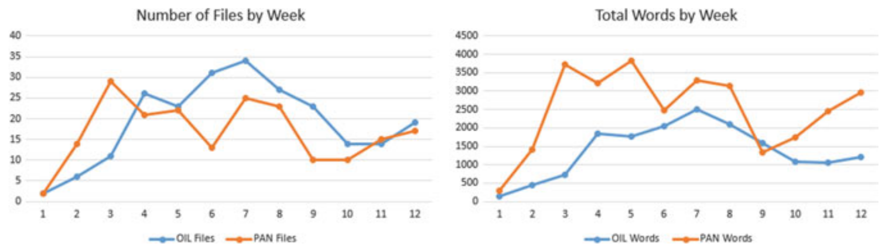


Fig. 13.1 Files per week and words per week graphs depicting each of the corpora

cross-corpus analysis), the effect of the differences in average word count per LTE were minimized.

13.3.2 Corpus Processing

After the corpora were aggregated, the LTEs were analyzed using simple term frequency techniques. All occurrences of non-stop nouns were counted in respect to both total number of occurrences as well as total number of unique LTEs in which they occurred. Many (though not all) of the expected (master) terms had a high enough score to be included in the weekly rankings. Additionally, some of the terms that were included in the weekly rankings were not among the master terms, allowing for unexpected terms to rise to levels of significance, adding depth and validation to the study.

A number of stop words were identified and ignored in the word frequency algorithm so that terms significant to the study (and relating to the two triggering events) could be aggregated and ranked. The list of stop words is provided in Appendix 13.6.

There were also a number of terms that were identified as being similar enough to be considered as equivalent. For example, the term “gas” in the context of a fuel shortage can be considered synonymous with the term “gasoline.” Although “gas” could also be used in a chemistry context, and even in an energy context (as with “natural gas”), a manual examination of the Oil Crisis corpus revealed that the primary use of the word “gas” was primarily used as a shortened form of the word “gasoline.” Thus, “gas” was tagged as an equivalent term for “gasoline.” Several other equivalent terms were determined in this same way. The inclusion of some terms but not other terms in the equivalent terms list had some effect on the final rankings for each corpus, so special care and much deliberation was made regarding the process of selecting equivalent terms. A list of all the equivalent terms used in this study are provided in Appendix 13.7.

13.3.3 Calculating Unique LTE Thresholds

LTEs tend to be influenced by trending dynamics, as a particularly volatile statement in one letter can ignite a number of letters over the next few days. Therefore, the trending dynamic of terms is just as important (if not more important) as the overall frequencies of the terms themselves. Some terms receive a high amount of attention for a brief period of time, and then are rarely written about again in subsequent weeks. However, the sheer magnitude of their frequencies over this brief period of time may skew the frequency data for all terms over the entire study. For this reason, a trending analysis tool was created for this paper based on a weekly ranking scheme that is popular among American sports, such as American football and basketball. These schemes rank teams according to their performance during that week and show upward and downward trending based on a comparison of their current week rankings with their ranks in the previous week. This scheme allows for an analysis of the significance of one term compared to another based on the number of LTEs that include a given term in a given week, as well as the overall frequency of that given term within the week's LTEs. The threshold for determining which terms are ranked week-by-week is called the Unique File Threshold.

The Unique File Threshold for each week was calculated by adding each week's number of files to the total number of files from the beginning of the study. The formula used for threshold calculation was int(totFiles*0.1) + 1, or roughly 10% of the total number of files accumulated up to and including that week (rounded down to the nearest integer) plus 1. For weeks where 25 words meeting the Unique Threshold for that week could not be generated by using the calculated Unique LTE Threshold value, a routine was provided to adjust the Unique File Threshold down incrementally by 1 until a ranking of the top 25 words were generated. Thus, some weeks have a Unique LTE Threshold one or two unique LTEs fewer than the original calculated value. The Unique LTE Threshold value for each week in each corpus is provided in Table 13.1.

13.3.4 Ranking the Top 25 Words

The top 25 words (aggregated values) for each week were copied to an Excel spreadsheet and aligned side-by-side so that the rankings of individual terms could be tracked from one week to the next in two dimensions: total times used and number of unique LTEs. The terms were scored by multiplying the total occurrences accumulated in all prior and current weeks by the total unique LTEs for the same period. Only those terms which met the minimum threshold of total unique LTEs for that week were eligible for ranking each week.

The weeks and unique LTE thresholds are provided for each of the corpora in Table 13.1.

Table 13.1 Date ranges, unique file thresholds, accumulated total files, and accumulated total words for each of the two study periods

Oil Crisis (OC) (cumulative from 10/23/1973)	OC unique file threshold	OC total files	OC total words	COVID Crisis (CC) (cumulative from 3/5/2020)	CC unique file threshold	CC total files	CC total words
To 10/29/1973	1	2	141	To 3/11/2020	1	2	282
To 11/5/1973	1	8	585	To 3/18/2020	2	16	1699
To 11/12/1973	2	19	1312	To 3/25/2020	5	45	5434
To 11/19/1973	5	45	3137	To 4/1/2020	7	66	8643
To 11/26/1973	7	68	4887	To 4/8/2020	9	88	12,468
To 12/3/1973	10	99	6935	To 4/15/2020	11	101	14,959
To 12/10/1973	13	133	9419	To 4/22/2020	13	126	18,262
To 12/17/1973	15	160	11,521	To 4/29/2020	15	149	21,390
To 12/24/1973	18	183	13,120	To 5/6/2020	16	159	22,723
To 12/31/1973	18	197	14,202	To 5/13/2020	17	169	24,458
To 1/7/1974	19	211	15,263	To 5/20/2020	19	184	26,917
To 1/14/1974	21	230	16,471	To 5/27/2020	21	201	29,882

Using the spreadsheet of aligned rankings, a shape or "curve" was generated following the ranking of significant specific terms over the 12-week study period.

The generated curve images are presented in the findings section.

13.4 Findings and Discussion

The findings generated through the weekly top 25 terms were produced to reveal a term's significance compared to all other terms used in the LTEs published in a given week. As the sports team rankings (the dynamics upon which these findings were based) attempt to reveal the strongest team that week which could, presumably, win a game that week if played against any other lower ranked team, the weekly top 25 term rankings in this study attempt to reveal the most significant terms used that week in rank order. For example, in a given week, if the term "gasoline" is ranked number one (which it was from Week 5 onwards), then it could be considered as a more significant term than the word “energy,” which encompasses more forms of energy than gasoline. In this model, if one were to walk into a diner and hear everyday ordinary people discussing current events over coffee, one would most likely hear more discussion about the effects of gasoline rationing and shortages than a discussion about the merits of oil as an energy source over solar power.

More than a tool to provide weekly comparisons, this ranking system also allows for visualizing trending data as well. As mentioned above, the term “gasoline” did not reach prominence until Week 5, where it stayed at the top of the rankings. One could interpret this “curve” produced by the term “gasoline” as a phenomenon where, at the beginning of the oil embargo, very few of the LTE writers anticipated the effect it would have on the economics of the gasoline market, both in supply and demand. As the weeks wore on, it became increasingly clear to average U.S. citizens that they were going to be affected most directly through the availability of gasoline, and efforts to conserve this commodity. This is reflected by the upward curve for the term “gasoline” (see Fig. 13.25).

Though there are other factors that may influence the visualized curve of a given term, the Unique File Threshold normalizes the frequencies of both the unique letters in which the terms appear that week, as well as the sheer frequency of the terms aggregated over all of the letters that week, so that the visualization of a term having high significance early in the study period and then “falling off” gradually (for example, retailer—OIL in Fig. 13.16) matches a manual reading of the LTEs over the weeks as well. A manual reader would come away from this task concluding that “retailers weren’t discussed much in the last few weeks of the study period.” However, the visualization is limited in that some LTEs are more powerfully written than others. There is no attempt in the visualization to quantify the level of impact that a specific letter may have on any individual reader or all readers collectively.

Generated rankings for both corpora are provided below. There are three columns included for each week’s rankings: the number of occurrences for that word, the word itself, and the number of unique files in which that term is found. A minimum of 25 words were ranked for each week, but words that had equal scores with the 25th ranked word were included as a “tie” in each week’s ranking. The only exception to this is week 1, which had several words beyond the 25 that had 1 occurrence in 1 file. In this case, words that persisted in other weeks were swapped into view with words that did not persist in other weeks.

The Oil Crisis rankings are revealed in Fig. 13.2

The COVID-19 Pandemic Crisis rankings are revealed in Fig. 13.3

In the rankings displayed in Figs. 13.2 and 13.3, the ranked terms were highlighted and those highlights then connected to form a visible “curve.” Of interest in these collections of ranked words for each corpus are three dimensions: the words that were present in the rankings (many that might be considered qualitatively significant were not included by the algorithm in rankings due to a low comparative score and/or having unique file counts lower than the minimum thresholds described above), the weekly ranking place of any specific word (based on a score calculated by the occurrences multiplied by the number of unique files compiled for that word in all

	Week	1		Week	2		Week	3		Week	4		Week	5		Week	6
2	mandate	2	11	sunday	5	14	conservat	11	29	conservat	21	50	gasoline	30	80	gasoline	47
2	conservat	2	7	energy	5	13	energy	10	29	energy	19	43	conservat	32	58	energy	43
3	retailer	1	5	conservat	4	12	sunday	6	26	gasoline	16	42	energy	29	55	conservat	41
3	oil	1	5	retailer	3	11	oil	6	26	light	16	29	light	17	43	light	25
3	arab	1	5	oil	3	7	retailer	5	15	retailer	10	31	school	13	34	retailer	20
2	open	1	4	open	3	7	gasoline	5	15	car	10	23	retailer	14	40	school	16
2	energy	1	3	panic	3	6	day	4	14	sunday	8	21	day	13	30	day	21
1	week	1	3	shortage	3	7	dst	3	12	governme	9	19	car	13	25	panic	24
1	war	1	3	food	2	5	open	4	15	school	7	22	president	11	30	president	17
1	usa	1	3	downtow	2	5	year	4	13	president	7	15	panic	15	26	car	19
1	three	1	3	day	2	5	light	4	10	people	9	19	sunday	11	23	shortage	20
1	sell	1	2	world	2	6	school	3	12	oil	7	17	open	12	26	hour	17
1	real	1	2	work	2	4	panic	4	11	day	7	19	hour	10	20	people	19
1	phantom	1	2	week	2	4	mandate	4	10	mandate	7	14	people	13	22	sunday	14
1	panic	1	2	usa	2	5	children	3	8	panic	8	14	year	11	23	oil	13
1	jets	1	2	people	2	7	hour	2	8	open	7	14	mandate	10	19	governme	15
1	israel	1	2	peace	2	4	people	3	8	little	7	12	shortage	11	20	open	14
1	industry	1	2	mandate	2	4	food	3	8	year	7	14	oil	9	17	year	14
1	fact	1	2	industry	2	3	usa	3	8	home	7	12	week	9	18	mandate	13
1	effort	1	2	husband	2	3	shortage	3	11	hour	5	12	governme	9	14	home	13
1	downtow	1	2	gasoline	2	3	industry	3	8	christmas	6	10	home	9	15	week	12
1	dollar	1	2	night	2	3	week	3	9	dst	5	12	dst	7	15	christmas	10
1	day	1	2	farmer	2	3	street	3	9	50	5	11	50	7	14	50	10
1	consumpt	1	3	president	1	3	home	3	7	week	6	9	little	8	11	omaha	11
1	close	1	3	arab	1	3	farmer	3	7	lot	6	9	close	7	11	way	10
1	blackmail	1	2	station	1	3	car	3	6	two	5	9	christmas	7	10	other	10
1	bad	1	2	shop	1	3	town	3	6	street	5						
1	almighty	1	2	ready	1	3	price	3	6	shortage	5						
1	aggressio	1	2	media	1												
			2	highway	1												

	Week	7		Week	8		Week	9		Week	10		Week	11		Week	12
107	gasoline	62	124	gasoline	73	139	gasoline	82	153	gasoline	91	161	gasoline	96	172	gasoline	105
65	energy	49	78	energy	60	88	energy	67	92	energy	70	104	energy	75	117	energy	83
63	conservat	47	69	conservat	53	75	conservat	59	80	conservat	63	85	conservat	67	96	conservat	74
50	light	29	59	light	32	63	light	35	65	light	37	63	day	44	78	oil	43
52	school	21	55	oil	29	55	day	38	58	day	40	70	oil	38	65	day	46
39	day	28	45	day	33	60	oil	32	65	oil	35	68	light	39	68	light	39
37	car	27	63	school	23	64	school	24	44	panic	38	55	car	39	60	car	43
38	oil	22	36	panic	31	42	panic	36	66	school	25	49	panic	42	53	panic	45
32	shortage	25	39	car	28	44	car	32	47	car	35	69	school	27	75	school	31
36	retailer	22	37	shortage	28	41	shortage	32	47	hour	31	50	hour	33	53	shortage	41
28	panic	26	41	retailer	25	43	hour	29	42	shortage	33	45	shortage	35	44	people	40
33	hour	22	37	hour	25	42	retailer	26	36	people	33	41	people	37	51	hour	34
34	president	21	31	people	28	32	people	29	42	retailer	26	41	president	27	43	president	29
26	people	24	34	president	21	36	president	23	34	usa	28	42	retailer	26	42	retailer	26
24	sunday	16	28	week	20	30	usa	26	34	governme	28	37	governme	29	37	governme	29
21	governme	17	26	governme	21	32	sunday	21	38	president	25	34	usa	28	34	usa	28
22	week	16	25	usa	21	29	governme	23	32	sunday	21	29	year	25	38	dst	23
20	year	16	28	sunday	18	33	truck	20	33	truck	20	34	truck	21	31	year	27
21	open	15	27	christmas	16	29	week	21	30	week	22	31	week	23	32	week	24
21	mandate	15	24	mandate	18	24	way	22	29	mandate	21	32	sunday	21	34	truck	21
17	usa	17	22	year	18	24	year	20	24	way	22	30	mandate	22	26	other	26
18	home	16	23	open	17	25	mandate	19	25	year	21	33	dst	19	27	way	25
22	christmas	13	20	way	19	24	open	18	23	home	21	26	way	24	32	sunday	21
16	omaha	14	20	home	18	21	home	19	24	open	18	25	home	23	28	omaha	24
14	way	13	18	little	15	20	omaha	18	21	omaha	19	22	other	22	30	mandate	22
13	other	13	15	other	15	18	other	18	22	little	18	22	omaha	20	26	home	24
									22	citizen	18						

Fig. 13.2 Oil Crisis word rankings by week

	Week	1		Week	2		Week	3		Week	4		Week	5		Week	6
9	covid	2	38	covid	14	77	covid	33	115	covid	50	152	covid	67	179	covid	76
3	burden	1	15	media	6	29	people	16	47	people	26	58	people	33	63	people	37
2	rate	1	8	work	6	24	work	17	43	media	23	49	media	29	52	nebraska	36
2	influenza	1	12	president	3	26	media	13	31	work	23	39	nebraska	28	51	media	31
2	health	1	7	omaha	5	18	other	15	25	other	22	34	other	29	39	other	33
2	father	1	6	health	5	17	panic	12	27	usa	20	34	work	26	47	president	21
2	efficacy	1	7	people	4	17	omaha	10	31	president	13	32	usa	24	36	work	27
2	care	1	5	public	5	13	health	12	23	panic	17	38	president	17	36	usa	27
1	year	1	6	panic	4	13	citizen	11	24	nebraska	16	31	omaha	20	35	health	25
1	words	1	6	influenza	4	13	usa	10	25	omaha	15	34	hospital	17	36	omaha	23
1	victim	1	4	citizen	3	20	president	6	20	citizen	17	27	panic	21	32	panic	25
1	story	1	4	conferenc	3	19	school	6	24	influenza	12	25	health	21	38	hospital	20
1	son	1	5	social	2	13	nebraska	8	18	health	15	24	citizen	20	35	governor	19
1	social	1	3	usa	3	10	public	10	16	public	15	29	governor	16	26	citizen	22
1	shot	1	3	other	3	11	influenza	7	16	year	12	22	worker	15	28	worker	19
1	season	1	3	decision	3	9	worker	8	17	day	9	20	year	15	28	care	18
1	safety	1	4	national	2	12	day	6	15	governor	10	24	influenza	12	24	public	20
1	rna	1	4	vaccine	2	9	situation	7	13	situation	11	20	home	14	25	home	19
1	recovery	1	4	question	2	9	home	7	20	school	7	17	public	16	23	year	18
1	quarantir	1	3	situation	2	8	life	7	14	hospital	10	21	day	12	25	risk	16
1	economy	1	3	risk	2	10	food	5	14	home	9	18	care	14	30	retailer	12
1	people	1	3	nebraska	2	7	local	6	12	worker	10	25	retailer	10	25	influenza	13
1	panic	1	3	mayor	2	7	care	6	11	life	10	24	school	9	23	day	14
1	omaha	1	3	leaders	2	6	hospital	6	12	local	9	15	situation	13	26	school	11
1	needn't	1	3	dr	2	6	first	6	12	risk	8	16	risk	11	21	restaurar	12
1	natural	1	3	county	2	6	economy	6	12	restaurar	8	15	local	10	16	world	15
1	mortality	1	3	human	2	6	due	6									
1	media	1															

	Week	7		Week	8		Week	9		Week	10		Week	11		Week	12
220	covid	93	256	covid	111	267	covid	118	281	covid	127	305	covid	138	336	covid	151
81	nebraska	50	96	people	55	106	nebraska	62	111	nebraska	66	127	nebraska	74	145	people	81
78	people	46	91	nebraska	53	101	people	57	108	people	60	123	people	70	131	nebraska	77
50	other	43	58	other	50	71	media	47	80	media	50	89	media	54	107	worker	55
57	media	36	65	media	42	60	other	51	65	other	55	93	worker	49	87	work	62
48	health	36	61	health	43	65	health	45	71	health	49	72	other	61	79	other	67
47	work	35	72	president	30	59	work	42	66	work	47	74	health	52	92	media	56
45	usa	34	55	work	39	64	worker	36	73	worker	39	72	work	52	83	health	56
54	president	24	60	worker	34	74	president	31	78	president	33	78	governor	45	85	governor	48
47	worker	27	52	usa	39	53	usa	40	54	usa	41	83	president	37	87	president	40
42	omaha	28	53	hospital	30	58	hospital	35	63	governor	35	64	hospital	40	63	usa	48
43	hospital	25	48	omaha	30	54	governor	32	59	hospital	36	58	usa	44	65	hospital	41
41	governor	23	49	governor	27	43	panic	34	45	panic	36	48	panic	38	61	care	38
44	retailer	21	41	panic	32	56	retailer	26	50	omaha	31	53	omaha	34	52	panic	42
34	panic	27	53	retailer	23	48	omaha	30	56	retailer	26	52	care	34	58	omaha	37
31	citizen	27	41	care	27	42	care	28	46	care	30	58	retailer	27	47	citizen	42
33	care	22	34	citizen	30	35	citizen	31	38	citizen	34	41	citizen	37	61	retailer	29
30	home	24	35	home	29	35	home	29	36	home	30	40	public	32	42	year	36
28	public	23	33	public	26	34	public	27	35	public	28	38	home	32	43	home	35
28	year	22	30	year	24	34	day	23	34	year	28	37	year	31	42	public	34
29	day	19	32	day	22	30	year	24	35	day	24	39	day	26	44	editorial	27
30	governme	15	27	world	23	33	editorial	20	36	editorial	21	39	editorial	24	41	day	28
25	risk	16	24	first	23	27	world	23	29	world	25	43	mask	20	49	mask	23
25	economy	16	28	editorial	18	24	first	23	40	mask	18	33	economy	22	38	economy	27
20	first	19	31	governme	16	29	economy	19	30	economy	20	29	world	25	34	world	29
28	school	13	27	economy	18	32	governme	17	25	first	24	35	governme	20	31	number	28
26	influenza	14										28	number	25			

Fig. 13.3 COVID-19 Pandemic Crisis word rankings by week

files up to that week), and the shape of the trending "curve" for a specific word (the change in ranking for that word from week to week).

The generated ranking curves from these two corpora were generally significant in 3 different ways:

- Some terms had sustaining curves in one corpus but were not present or had minimal presence in the other corpus
- Some terms had sustaining curves in both corpora, allowing them to be compared from between the corpora
- A term with a sustaining curve in one corpus had a similar curve to that of a different term in the other corpus.

These three sets of findings are presented and discussed separately below. It should be noted that the findings presented below are those which revealed themselves primarily through their visual trending curves. Additional insights may be gleaned from focusing on terms which did not have visual trending curves that drew our attention, but still may reveal some significant insights into these two crises.

It should also be noted that the interpretation of the role of such a term as "citizen" has been undertaken solely from how that term was used in a manual investigation of the LTE texts in which the term has been included. Other researchers may have different interpretations, which may be just as valid as the interpretations provided here. The interpretations are provided more as a model for discussion than an absolute. A sampling of the LTEs can be accessed directly by readers at http://www.kokensparger.com/brian/exploration/LTEs/, so that interested readers may see the terms in the context of selected LTEs in which they appear. This will allow readers to follow our path in reading through the letters and interpreting the roles of the terms therein.

13.4.1 Persistent Curves in One Corpus Only

There were seven terms with persistent curves in one corpus that were not present or had minimal presence in the other corpus: "citizen" (Fig. 13.4), "government" (Fig. 13.5), "media" (Fig. 13.6), "omaha" (Fig. 13.7), "other" (Fig. 13.8), "work" (Fig. 13.9) and "world" (Fig. 13.10).

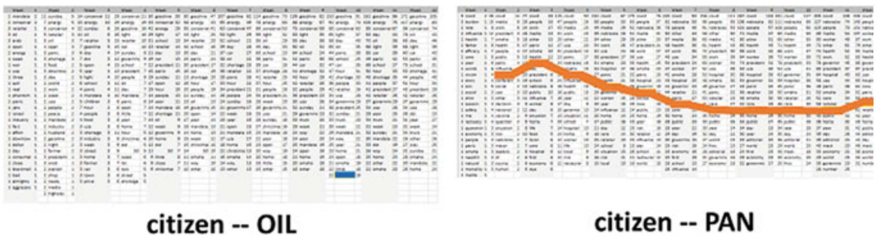


Fig. 13.4 Ranking shapes (curves) for the term "citizen" as it occurs in both the OIL and PAN corpora

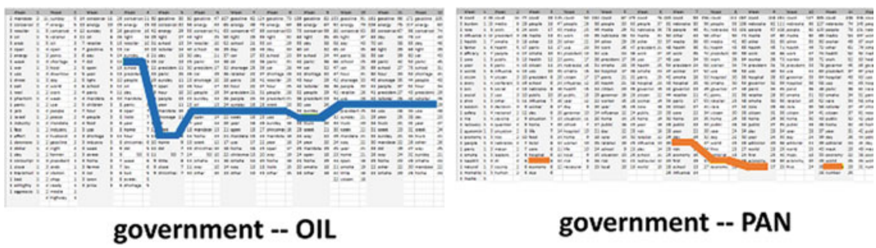


Fig. 13.5 Ranking shapes (curves) for the term "government" as it occurs in both the OIL and PAN corpora

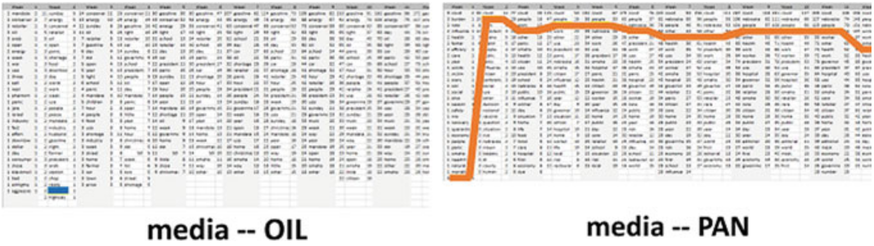


Fig. 13.6 Ranking shapes (curves) for the term "media" as it occurs in both the OIL and PAN corpora

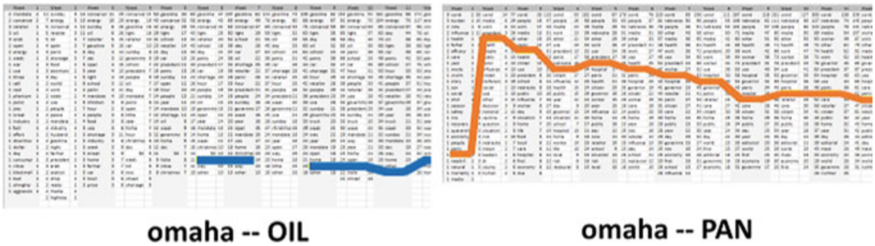


Fig. 13.7 Ranking shapes (curves) for the term "omaha" as it occurs in both the OIL and PAN corpora

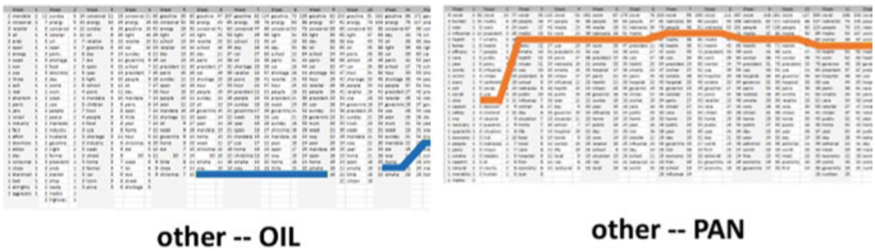


Fig. 13.8 Ranking shapes (curves) for the term "other" as it occurs in both the OIL and PAN corpora

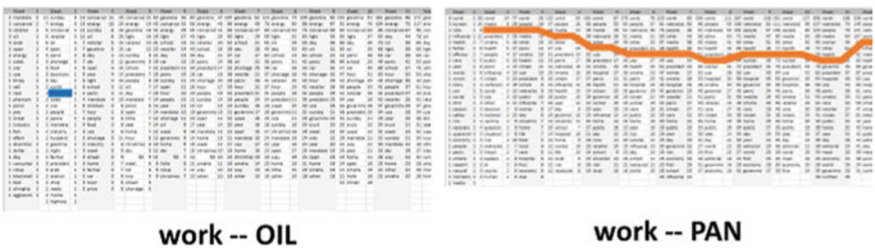


Fig. 13.9 Ranking shapes (curves) for the term "work" as it occurs in both the OIL and PAN corpora

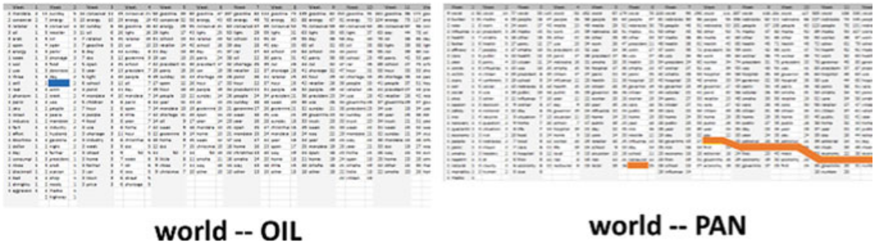


Fig. 13.10 Ranking shapes (curves) for the term "world" as it occurs in both the OIL and PAN corpora

In Fig. 13.4, the term "citizen" was sustained within the LTEs consistently within the COVID Crisis (PAN) corpus but not consistently within the Oil Crisis (OIL) corpus over the study period. As a collective term for "American," the term "citizen" comes up much more prominently during the COVID Pandemic Crisis than the Oil Crisis. Could it be due to the rise in nationalism during the Trump administration? As a proposition, this phenomenon is also supported in other terms that were more significant in the PAN corpus, such as "media" and "Omaha" (i.e., focus on local government's failings over those of the federal government).

In Fig. 13.5, the term "government" (with its equivalent terms of "washington" and "lincoln"—the state capitol of Nebraska) was sustained within the LTEs consistently within the Oil Crisis (OIL) corpus but not consistently within the COVID Crisis (PAN) corpus over the study period. The term "government" was more prominent as a collective term during the Oil Crisis than during the COVID Crisis. In LTEs, a term like "I think Washington should... " can be interpreted as the same statement as "I think the government should... " The term "government" was used as a collective term for all levels of government much more during the Oil Crisis, but much less during the COVID crisis. This could simply be a difference between the 1970s, where citizens were more likely to refer to the Federal government as "Washington" and the State government as "Lincoln," and the more current times where citizens refer more to the government by using the names of the leaders (a.k.a., "Trump" and "Ricketts", respectively).

In Fig. 13.6, the term "media" (with its equivalent terms of "newspaper", "internet", "world-herald", "television", "tv", "facebook", "twitter", "tweet", "press", "publish", "paper" and "news") was sustained within the LTEs consistently within the COVID Crisis (PAN) corpus but not consistently within the Oil Crisis (OIL) corpus over the study period. These curves are similar to the curves for "citizen"—"media" (and all its equivalent terms, like "press" and "tv" and "newspapers") has a much higher place in the discussions during the COVID Crisis, perhaps in connection to the Trump Administration's attack on the media during his administration.

In Fig. 13.7, the term "omaha" (with its equivalent term of "city") was sustained within the LTEs consistently within the COVID Crisis (PAN) corpus but not consistently within the Oil Crisis (OIL) corpus over the study period. As "government" as a collective term had higher significance during the Oil Crisis, the more specific

"Omaha" had higher significance during the COVID Crisis, as there was a heightened sensitivity to conflict among and between the various levels of government. Some of this may be due to the combative nature between the federal government, the state government (Nebraska), and the local government (Omaha).

In Fig. 13.8, the term "other" was sustained within the LTEs consistently within the COVID Crisis (PAN) corpus but not consistently within the Oil Crisis (OIL) corpus over the study period. The term "other" is interesting in that it suggests that alternatives were somehow considered by the LTE writers, and upon examination of the term within the LTEs, there were several contexts in which the term was used. This implicit investigation of alternatives perhaps is the opposite of "duty"—could the sense of doing one's duty have degraded between the 1970s and the 2020s?

In Fig. 13.9, the term "work" (with its equivalent terms of "job" and "employment") was sustained within the LTEs consistently within the COVID Crisis (PAN) corpus but not consistently within the Oil Crisis (OIL) corpus over the study period. The term "work" has much greater significance during the COVID Crisis, as the economy during that crisis was almost instantly closed down, causing a huge upswing in unemployment. Work became an issue among LTE writers and, though it leveled off a bit, persisted in the top ten of rankings from its entrances in week 2 through the end of the study period.

In Fig. 13.10, the term "world" was sustained within the LTEs consistently within the COVID Crisis (PAN) corpus but not consistently within the Oil Crisis (OIL) corpus over the study period. Although the Oil Crisis was, in a real sense, a "world" problem, the discussions among the LTE writers largely ignored the global aspect of this crisis to focus on the national and local aspects. During the COVID Crisis, "world" as a term was much more significant, though not largely occurring in the LTEs until later in the study period. This may partially due to the uptick of discussion about the World Health Organization, and President Trump's squabble with it.

13.4.2 Persistent Curves in Both Corpora

Some terms had persistent curves in both corpora, revealing consideration that they are present in the LTE discussions around any crisis of national scale or larger. The terms "day" (Fig. 13.11), "home" (Fig. 13.12), "panic" (Fig. 13.13), "people" (Fig. 13.14), "president" (Fig. 13.15), "retailer" (Fig. 13.16), "school" (Fig. 13.17), "usa" (Fig. 13.18) and "year" (Fig. 13.19) all were used liberally and often in both corpora. They are briefly discussed under their visualized curves below.

In Fig. 13.11, the term "day" was sustained within the LTEs consistently over both corpora during the study period. The usage of the term varied from historic accounts (e.g., "in my day") to accounts of struggle (e.g., "it's difficult to get through the day") and other sentiments that crises are lived day-by-day. The higher significance and more persistent curve during the Oil Crisis may be largely due to the fact that the word "day" was often tied into the concept of daylight savings time (dst), which was one of the top terms mentioned among the LTEs in that corpus. Daylight savings

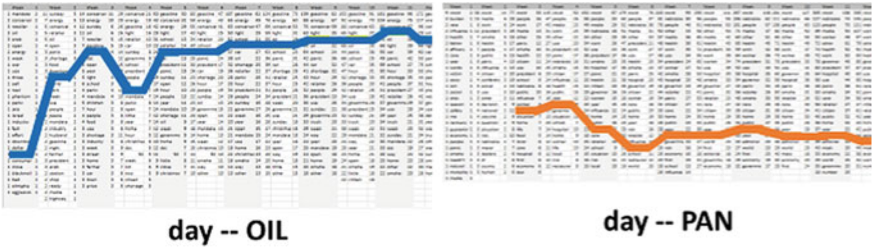


Fig. 13.11 Ranking shapes (curves) for the term "day" as it occurs in both the OIL and PAN corpora

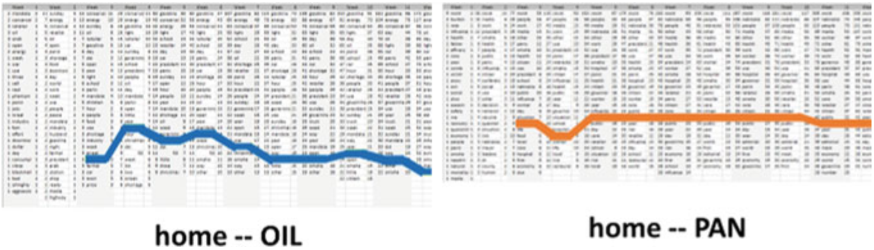


Fig. 13.12 Ranking shapes (curves) for the term "home" as it occurs in both the OIL and PAN corpora

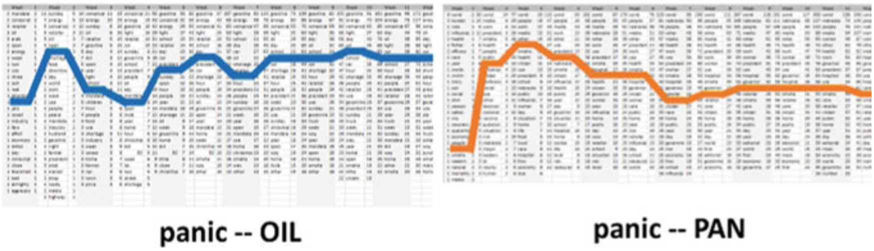


Fig. 13.13 Ranking shapes (curves) for the term "panic" as it occurs in both the OIL and PAN corpora

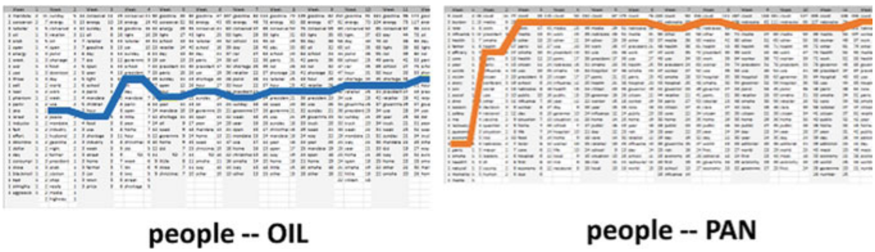


Fig. 13.14 Ranking shapes (curves) for the term "people" as it occurs in both the OIL and PAN corpora

Fig. 13.15 Ranking shapes (curves) for the term “president” as it occurs in both the OIL and PAN corpora

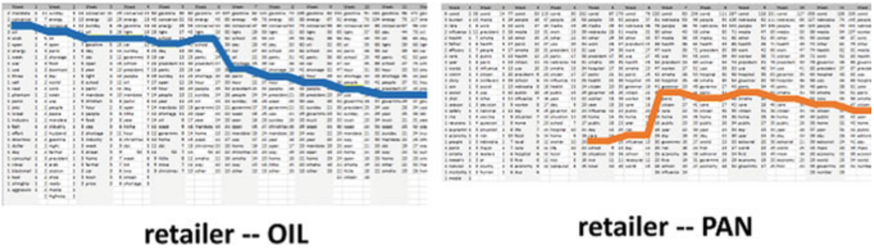


Fig. 13.16 Ranking shapes (curves) for the term “retailer” as it occurs in both the OIL and PAN corpora

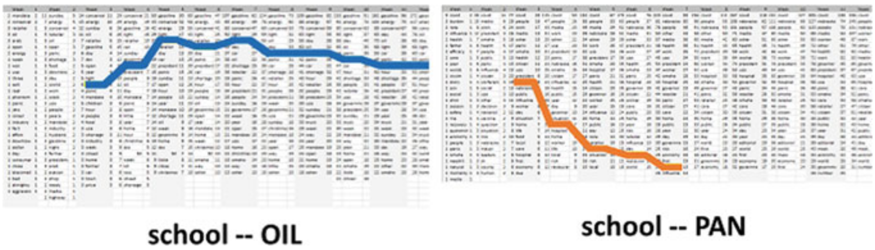


Fig. 13.17 Ranking shapes (curves) for the term “school” as it occurs in both the OIL and PAN corpora

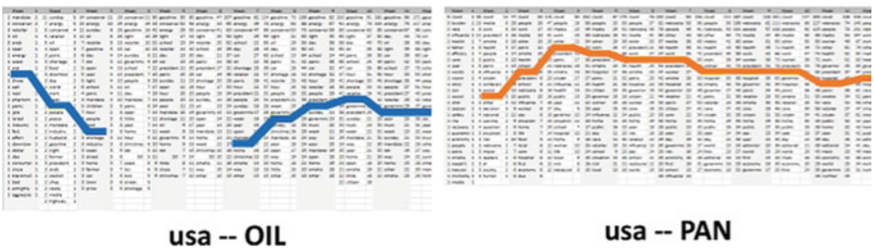


Fig. 13.18 Ranking shapes (curves) for the term “usa” as it occurs in both the OIL and PAN corpora

Fig. 13.19 Ranking shapes (curves) for the term "year" as it occurs in both the OIL and PAN corpora

time (dst)—turning clocks ahead for an hour in the spring and back again in the fall—was established in the United States in 1918 as a way to save fuel (by having more sunlight in the evenings for recreation and leisure). During the oil crisis, dst was proposed as a fuel saving measure but was unpopular because it meant that children would need to rise and walk to school in the dark. The process of finding and aggregating equivalent terms for dst undoubtedly caught some uses of the word "day" in this way, but not all of them.

In Fig. 13.12, the term "home" was sustained within the LTEs consistently over both corpora during the study period. Homes are significant topics for LTEs during crises because they are, literally, where people live. Although the curve for "home" during the Oil Crisis study period was a bit more erratic than that for the COVID Crisis, the relative scores for the word in both corpora were roughly the same. From these curves, it could be interpreted that the term "home" had equal usage during both crises and might be expected to show up similarly in any future crisis.

In Fig. 13.13, the term "panic" (and its equivalent terms, including "hysteria", "crisis", "ordeal" and "emergency") was sustained within the LTEs consistently over both corpora during the study period. As both corpora were collected following "crisis" trigger articles in the same newspaper, it is expected for this term to show up significantly in both corpora, and at generally equivalent relative placings in their weekly rankings. If nothing else, this term provides a bit of validation regarding the collection and analysis methodology of this study.

In Fig. 13.14, the term "people" was sustained within the LTEs consistently over both corpora during the study period. That any crisis is centered on people, and therefore the term "people" entered the LTE discussions regularly, is no surprise. The genuinely higher ranking in the COVID Crisis (PAN) corpus may be generated by a higher-than-normal trend of using the term "people" in emotional statements, such as "Come on people, wake up!".

In Fig. 13.15, the term "president" (and its equivalent terms, including "nixon's", "trump's", "richard r. nixon", "donald j. trump", "richard nixon", "donald trump", "the administration", "the administration's", "potus", "nixon", and "trump") was sustained within the LTEs consistently over both corpora during the study period. Although the term appears and then drops out of the rankings and re-appears at a much higher rank in the Oil Crisis (OIL) corpus, from that point on the general rank of the term is consistent between both corpora. As both crises were national-level crises, it is expected that the President of the United States would be discussed in many of the LTEs during both study periods. These curves suggest that in any national-level crises, the American populace looks towards the President for leadership and direction on how to deal with the crises.

In Fig. 13.16, the term "retailer" (and its equivalent terms, "including", "retailers", "wholesaler", "wholesalers", "superstore", "superstores", "malls", "mall", "shopping centers", "shopping center", "supermarkets", "supermarket", "big box stores", "big box store", "stores", and "store"), was sustained within the LTEs consistently over both corpora during the study period. In the Oil Crisis (OIL) corpus, the term "retailer" began at a high ranking and dropped gradually to midrange by the end of the study period. In the COVID Crisis (PAN), the term also dropped gradually, though it did not enter the rankings until halfway through the study period. This indicates that the OIL Crisis discussion began immediately after the trigger article was written about retailers reducing hours, and perhaps closing for entire days like Sundays, as this was one of the earliest suggested energy conservation measures from the president. During the COVID Crisis, the impact of retailers was not felt until individual cities (such as New York City) began locking down businesses, with Nebraska cities following suit shortly thereafter.

In Fig. 13.17, the term "school" (and its equivalent terms, including "schools", "classroom", "classrooms", "classes", "schoolroom", and "class") was sustained within the LTEs consistently over both corpora during the study period. School is a place, like home, where a crisis has an immediate impact. In both corpora, discussions in the LTEs around the term "school" began in week 3. In the Oil Crisis (OIL), the discussion continued, and the term was used frequently for the remainder of the study period. This was most likely due to two factors: As the Oil Crisis study period covered late autumn into winter, school was ongoing (whereas the COVID Crisis (PAN) occurred in late spring, when schools were poised to end their sessions for the summer anyway), and the daylight savings (dst) discussion continued to involve use of the term due to the emotional idea of children walking to school in the dark.

In Fig. 13.18, the term "usa" (and its equivalent terms, including "united states", "u.s.a.", "u.s." and "america") was sustained within the LTEs consistently over both corpora during the study period. As with the term "citizen" (which has as an equivalent term "American"), the higher ranking of the term during the COVID Crisis (PAN) could be attributed to the rise in nationalism under the Trump administration. The disappearance of the term in the Oil Crisis (OIL) rankings is a bit of a mystery, though the appearance of the term "50" (most likely referring to the 50 mph suggested speed limit) may have supplanted feelings of national pride during that time.

In Fig. 13.19, the term "year" was sustained within the LTEs consistently over both corpora during the study period. In terms of time frames, as "day" is significant during crises, a "year" is also significant, as these two periods of time are the most prevalent ones used in LTE discussions (over terms like "week" and "month" and even "decade"). It might be construed that the normal unit of time for discussions in the LTEs was the day, but the long-range time frame is the year, with most of the writers not willing to look much further ahead than that period of time.

13.4.3 Different Terms with Similar Curves

In comparing curves from terms in both corpora, a number of different terms seemed to have the same dynamics in their respective corpora. These similarities are of special interest in this study, because they may indicate that some terms "behave" similarly, or perhaps have similar roles in their respective corpora. This is especially intriguing when the same term in the other corpus has different dynamics. What follows are the findings of terms in this category.

In Fig. 13.20, the term "energy" was sustained within the Oil Crisis (OIL) corpus and "covid" (with its equivalent terms of "coronavirus", "virus", "pandemic", "covid-19", "epidemic", and "covid's") was sustained within the COVID Crisis (PAN) corpus with similar curves during the study period. Both of these terms have consistently high rankings in their respective corpora, as was expected, since these terms were used to choose specific LTEs to include in the corpora. Therefore, these terms are fundamental terms in the corpora, and would be expected to be the top term. The fact that "energy" is not the top term in the Oil Crisis corpus is surprising—though it was a discriminating term, it was not the only term, with "gasoline" (and its equivalent terms) occurring more often and in more unique LTEs.

In Fig. 13.21, the term "shortage" was sustained within the Oil Crisis (OIL) corpus and "worker" (with its equivalent terms of "employees", "workers", "professionals", "employee", and "professional") was sustained within the COVID Crisis (PAN) corpus with similar curves during the study period. The idea of a shortage—especially a fuel shortage in a society that highly valued its automobiles—had a high impact early in the study period (week 2), then fell sharply in the rankings, and then rose equally sharply in the middle weeks, leveling off in the final weeks. This seems to indicate that the early high rankings were tied to fear of a shortage, as such fears generally subsides when there was not an ongoing visible threat. Then later, as fuel shortages began to show, concern rose again, but based on the reality, not the fear. In regard to the term "worker", some of the same dynamics appear, though the fall and rise parts of the curve are not as pronounced. Could the "shortage" of the COVID Crisis be the worker, especially the health care worker? There are also similar curves

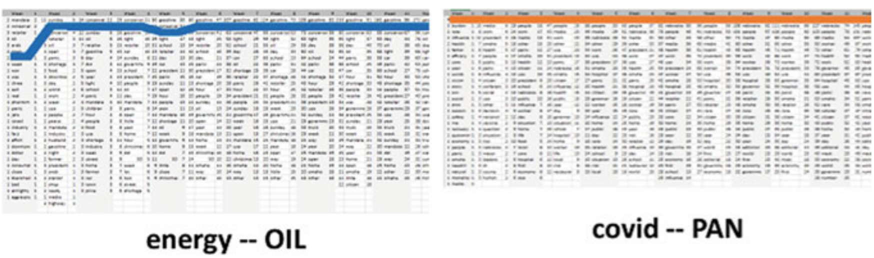


Fig. 13.20 Ranking shapes (curves) for the term "energy" as it occurs in the OIL corpus and "covid" as it occurs in the PAN corpus

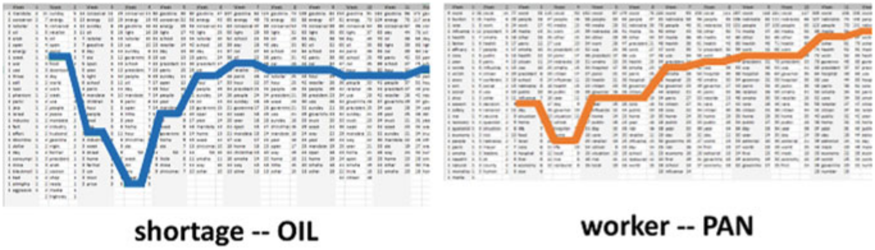


Fig. 13.21 Ranking shapes (curves) for the term “shortage” as it occurs in the OIL corpus and “worker” as it occurs in the PAN corpus

(discussed below) in the “oil” term (in the Oil Crisis) and the “health” term (in the COVID Crisis), adding support to this equivalency of oil/shortage and health/worker.

In Fig. 13.22, the term “car” was sustained within the Oil Crisis (OIL) corpus and “hospital” (with its equivalent terms of “medical”, “clinic”, “hospitals”, “clinics” and “medical facility”) was sustained within the COVID Crisis (PAN) corpus with similar curves during the study period. Both of these curves begin relatively low in ranking, have a sharp gain in ranking, and then level off. In terms of an analogy, a car is related to the Oil Crisis (directly affected in a way that is perceivable to most of the population) as a hospital is related to the COVID Crisis (again, directly affected in a way that is perceivable to most of the population). They are both terms for places where the LTE writers might directly “feel the pain” of the crises. Many of the LTEs focused on differing aspects of cars—when and how people should drive them and how much gasoline they should be allowed to use during the Oil Crisis. During the COVID Crisis, the letter writers “felt the pain” of the pandemic in the hospitals—where they focused on when and how people should arrive at a hospital for treatment, and which ailments were serious enough to use hospital resources that could otherwise be used for COVID patients. These terms provide further evidence that different words that define the parameters of different crises may “behave the same” or fulfill the same roles in the LTE discussions about the crisis.

In Fig. 13.23, the term “conservation” (with its equivalent terms of “saved“, “savings”, “saving”, “reduced”, “reduction”, “cut down”, “turn down”, “conserved”,

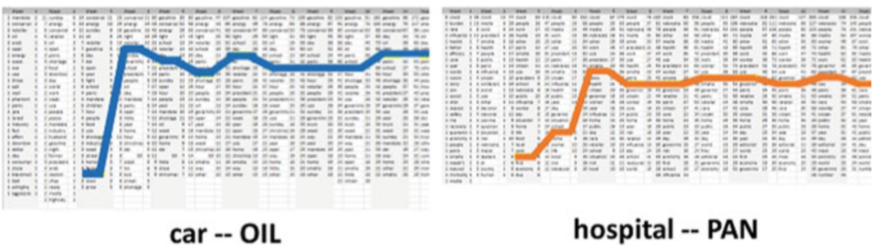


Fig. 13.22 Ranking shapes (curves) for the term “car” as it occurs in the OIL corpus and “hospital” as it occurs in the PAN corpus

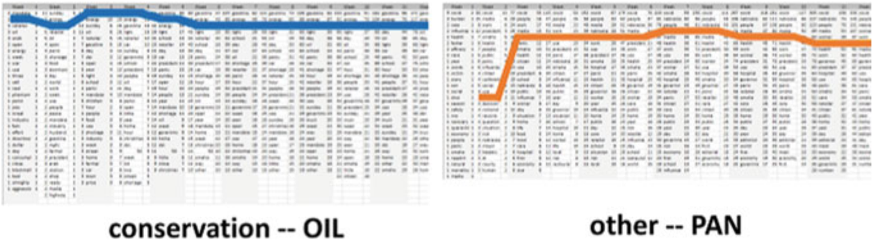


Fig. 13.23 Ranking shapes (curves) for the term "conservation" as it occurs in the OIL corpus and "other" as it occurs in the PAN corpus

"conserves", "conserve", "turn off", "saves", "reduce" and "save") was sustained within the Oil Crisis (OIL) corpus and "other" was sustained within the COVID Crisis (PAN) corpus with similar curves during the study period. Both terms are relatively surprising high-ranking terms in their corpora, particularly the term "other" in the COVID Crisis corpus. Both of these terms are tied to the idea of alternatives. In the Oil Crisis, "conservation" suggests altering common behavior, and advocates for a culture change. This played out as attacks by LTE writers on Sunday leisure driving and displaying holiday lights in stores and at homes. As there is a strong emotional attachment to these practices, there was an equally strong discussion of the call to give them up by governmental agencies. Though it is much more subtle, the term "other" had similar dynamics in the COVID Crisis. The behavioral changes forced upon the average citizen during this crisis were as intense as the conservation approaches required by the Oil Crisis. Citizens were asked to avoid gathering in large groups, to not attend sporting events, and to stay home from church. Some of the LTEs considered other ways to protect the citizens from the COVID pandemic than the recommendations provided by the government. As such, it could be interpreted that the two terms "behave" similarly in their uses in each of their respective corpora. Readers may consider this further by viewing the selected LTEs available on the website indicated above.

In Fig. 13.24, the term "mandate" (with its equivalent terms of "requirements", "requirement", "laws", "law", "mandates", "limits", "limit", "sanctions" and "sanction") was sustained within the Oil Crisis (OIL) corpus and "public" was sustained within the COVID Crisis (PAN) corpus with similar curves during the study period. The curves between these two terms are eerily similar, starting with high rankings, falling off rather quickly, and then leveling off at mid-low levels. These terms equate in the sense that a public manifestation of the law is a mandate, and a mandate moves a private behavior or practice into the public for examination.

In Fig. 13.25, the term "gasoline" (with its equivalent terms of "fuel", "petroleum", "ethel", "gasahol" and "gas") was sustained within the Oil Crisis (OIL) corpus and "media" (with its equivalent terms of "newspaper", "internet", "world-herald", "television", "tv", "facebook", "twitter", "tweet", "press", "publish", "paper" and "news") was sustained within the COVID Crisis (PAN) corpus with similar curves during the study period. Both of these curves begin rather lowly-ranked, and have a huge gain

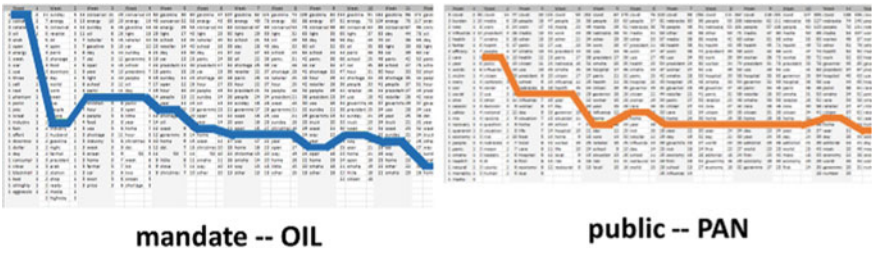


Fig. 13.24 Ranking shapes (curves) for the term "mandate" as it occurs in the OIL corpus and "public" as it occurs in the PAN corpus

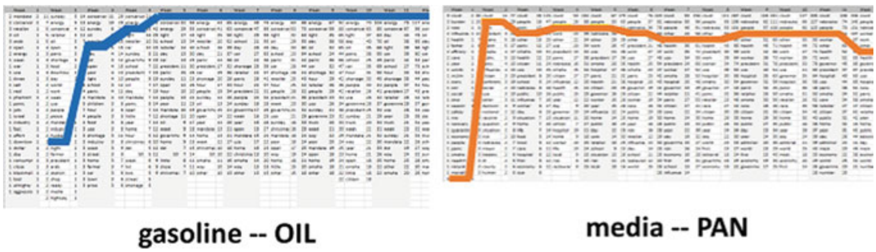


Fig. 13.25 Ranking shapes (curves) for the term "gasoline" as it occurs in the OIL corpus and "media" as it occurs in the PAN corpus

over a week (with "gasoline" taking another large step), then both leveling off at a high ranking for the remainder of the study period ("media" taking a step down at the end of this period). Such a large rise and a high ranking among these terms suggest that they are both integral to the crisis and are also emotionally volatile. Both of these terms represent "inflammatory emotional battleground of the crisis," which represents the intense emotional discussion surrounding the management of the gasoline supply during the Oil Crisis and management of the media during the beginning of the COVID Crisis, especially during the Trump Administration.

In Fig. 13.26, the term "open" was sustained within the Oil Crisis (OIL) corpus and "influenza" (with its equivalent terms of "flu virus", "flu epidemic", "flu pandemic" and "flu") was sustained within the COVID Crisis (PAN) corpus with similar curves during the study period. Both curves begin with high rankings and drop steadily until they are eventually out of the rankings at or just past the middle weeks in the study. The term "influenza" is perhaps the easier term to interpret; early in the COVID Crisis, direct comparisons were made between the COVID and Influenza, with the hope that the COVID virus would only affect the American populace as the influenza virus does now, and that it would not affect the populace as the 1918 Influenza did (killing millions). Within these dynamics, the term "open" is a little more difficult to interpret. The early discussions during the Oil Crisis centered around whether or not stores should remain open, as the chief controversy during those weeks. The

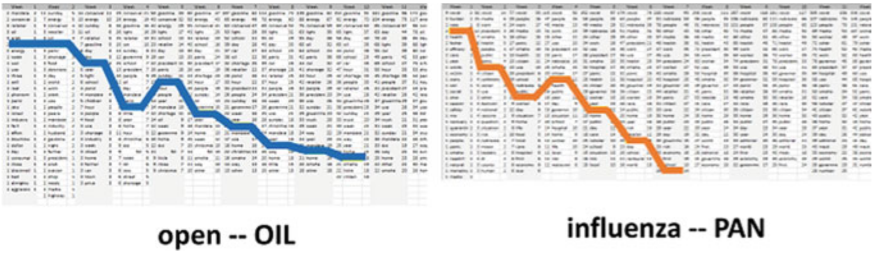


Fig. 13.26 Ranking shapes (curves) for the term “open” as it occurs in the OIL corpus and “influenza” as it occurs in the PAN corpus

discussions reduced as other issues arose, such as daylight savings time, the 50 mph speed limit, and gas rationing, to replace this early term of focus.

In Fig. 13.27, the term “oil” was sustained within the Oil Crisis (OIL) corpus and “health” sustained within the COVID Crisis (PAN) corpus with similar curves during the study period. The “oil” curve begins with high rankings in the Oil Crisis corpus, drops solidly for 3 weeks, rises just as solidly for 3 weeks, and then levels off at a relatively high ranking. The “health” curve is similar to the “oil” curve, including the 6 week drop and rise dynamic, except the rankings do not drop as far. This shows an equivalent role for these terms in their respective corpora. It could be that these terms reveal the general real “battleground” for their crises: the energy crisis is really about oil as the kingpin of energy, and the COVID crisis is about health as the key battleground for the pandemic crisis. As mentioned above, these roles are similar in behavior to the connected roles of “shortage” and “worker”.

In Fig. 13.28, the term “truck” was sustained within the Oil Crisis (OIL) corpus and “mask” (with its equivalent terms of “masks”, “face covering”, “face coverings”) was sustained within the COVID Crisis (PAN) corpus with similar curves during the study period. Both curves start late in the study period and reveal a relatively consistent presence in LTE discussions after their initial appearances. These dynamics reveal late but virulent issues. During the Oil Crisis, there was a brief number of truck blockades of the Interstate highways, especially in the states of Ohio and Pennsylvania. Public outcry arose and permeated the Public Pulse for a

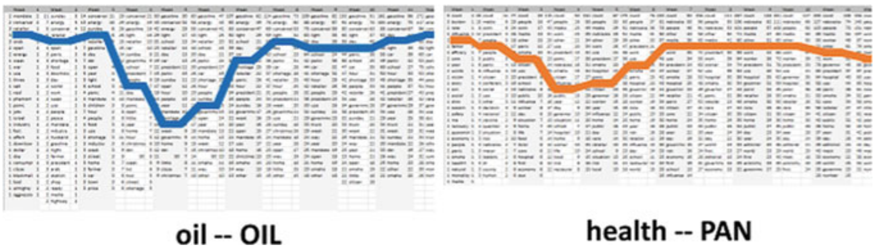


Fig. 13.27 Ranking shapes (curves) for the term “oil” as it occurs in the OIL corpus and “health” as it occurs in the PAN corpus

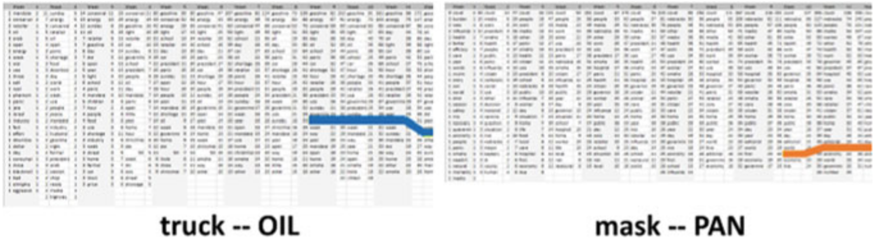


Fig. 13.28 Ranking shapes (curves) for the term "truck" as it occurs in the OIL corpus and "mask" as it occurs in the PAN corpus

few weeks. Similarly, during the COVID Crisis, the prospect of requiring that masks be worn in public spaces was communicated, followed by a public outcry of positive and negative responses.

13.4.4 Additional Interesting Curves Produced in the Study

Some additional curves, and aggregated curves, deserve some discussion as findings for this study. They include the terms "sunday" (Fig. 13.29), "mayor/governor/president" (Fig. 13.30), and "dst/day/light" (Fig. 13.31).

One of the earliest suggested conservation measures was to close retail stores and gas stations on Sundays. This discussion vaulted the term "Sunday" to the top of the rankings in week 2. As other conservation measures were introduced, the controversial Sunday closing enforcement gradually subsides, until Christmas week gives the word a good lift as shoppers contemplate if their favorite stores will be closed on Dec. 23rd, the Sunday before Christmas.

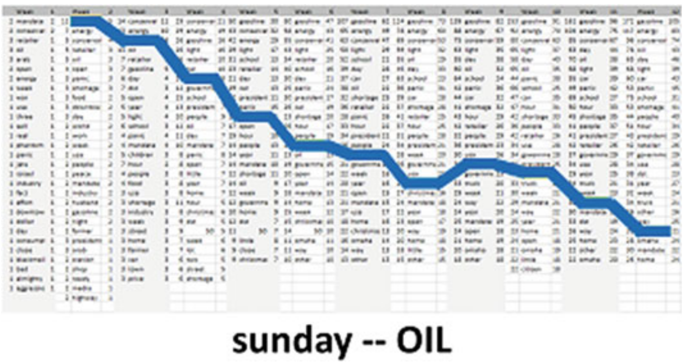


Fig. 13.29 Ranking shapes (curves) for the term "sunday" as it occurs in the OIL corpus

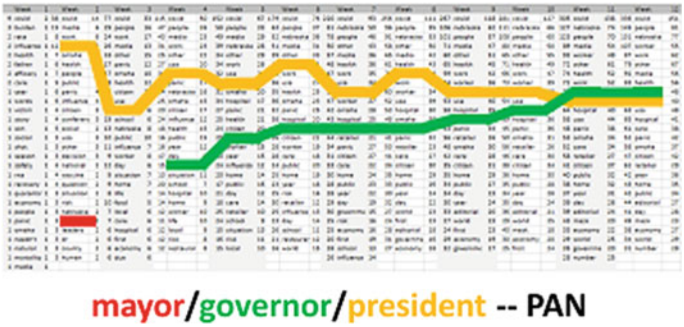


Fig. 13.30 Ranking shapes (curves) for the terms "mayor", "governor", and "president" as they occur in the PAN corpus

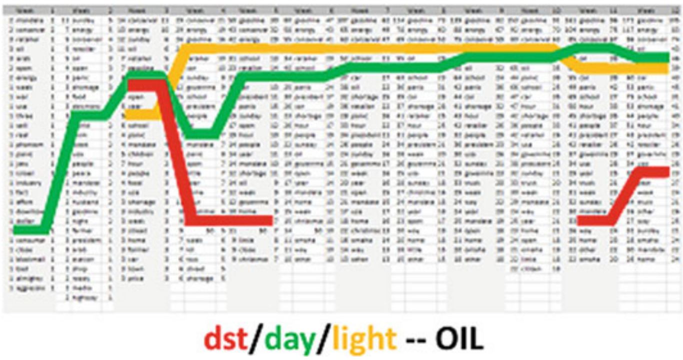


Fig. 13.31 Ranking shapes (curves) for the terms "dst", "day", and "light" as they occur in the OIL corpus

During the COVID Pandemic the stark difference between LTEs mentioning the president and governor disappears as the fight between President Trump and the governors reflected the struggle in the state of Nebraska. Although the mayor of Omaha was instrumental in proposing several restrictions in the city, she was only moderately mentioned, and not mentioned nearly as much as the president and governor.

During the Oil Crisis, a controversial conservation measure suggesting that daylight savings time (dst) be implemented was introduced and discussed frequently in the LTEs. The term dst is not always captured by the system's equivalent terms, as the writers used several euphemisms to refer to the measure, such as a "shorter day" and "no light in the morning," so these curves are most likely a better representation of dst.

13.5 Conclusions

During a crisis, specific places such as “home”, “school”, “retailer”, and “work” garner special attention in the LTEs, while time frames of “day” and “year” also garner special attention.

LTE writers in the United States will verbalize their crisis in a number of ways and will look to the President for leadership and guidance. How the President provides this leadership will be reflected in either a sense of nationalism (“citizen”, “usa”, “America”, “American”) or a relative objectification of it (“government”).

The media’s role (collectively) will be evaluated and judged by the LTE writers, and—in times of nationalism—will generally reflect how it is viewed by the President and will thus be discussed significantly in the LTEs. During other times, it will be nonexistent as a term.

The use of the term “other” signals a level of dissatisfaction of the LTE writers with suggestions and guidelines, as well as a judgment borne out by personal observations. It may also signal a feeling of isolation versus camaraderie.

Several different terms between the two corpora parallel one another in their functional dynamics but adhere to their counterparts within their corpora. Table 13.2 shows these parallels.

As the terms in Table 13.2 “behave” similarly or perform similar roles in the discussions during their respective crises, it could be expected that, with the arrival of a new national-level crisis of an entirely different source, that terms with similar curves may have significant rankings in the corpus produced by its LTEs. For example, if a national-level threat arose from an extraordinary number of disappearing teenagers, and an article was printed in the *Omaha World-Herald* reporting this trend and suggesting that everyone become more aware of their teenagers’ activities over the next few days, we could expect a number of terms with similar usage in the LTEs to follow the report (especially if there were a number of subsequent articles on the ongoing crisis).

Table 13.2 Functional dynamics of different terms with similar curves in their respective corpora

Oil corpus term	COVID corpus TERM	Functional dynamic
Energy	COVID	Term for selection
Shortage	Worker	Focus of fear
Car	Hospital	Places where citizens feel the “pain” of the crisis
Conservation	Other	Discussion around recommended alternatives to common behaviors
Mandate	Public	Externalization of the crisis
Gasoline	Media	Inflammatory emotional battleground of the crisis
Open	Influenza	Early vector for denial
Oil	Health	Big-picture battleground of the crisis
Truck	Mask	Later emotionally inflammatory trigger term

Applying the examples of the list in Table 13.2 to this hypothetical crisis, words such as "abduction", "runaway", "wellbeing", "curfew", "group", "family", "friend", "stranger", "trust", and perhaps later in the study period, an emotional inflammatory trigger term like "lockup" or "surveillance" might arise in the rankings by a similar analysis performed on the corpus of LTEs generated over the 12 weeks following the hypothetical trigger article.

Appendices

13.6 Appendix

Stop words for this study include:

stopWords = ["the","and","to", "a", "of", "i", "in", "my", "you", "his", "that", "is","he","it", "not", "ha", "\"can't\", "for", "on", "be", "me", "with", "was", "we", "as", "from", "but", "are", "when", "like", "our", "this", "am", "say", "her", "will", "so", "our", "or", "into", "if", "they", "out", "by", "one", "who", "what", "let", "have", "has", "more", "at", "all", "too", "she", "had", "i'm", "its", "an", "time", "then", "some", "how", "him", "*", "us", "them", "own", "only", "just", "here", "can", "&", "were", "yet", "than", "it's", "do", "your", "these", "their", "there", "would", "where", "after", "because", "every", "even", "each", "through", "over", "came", "always", "about", "up", "down", "tried", "now", "must", "made", "know", "form", "don't", "cannot", "beneath", "before", "become", "no", "without", "still", "back", "across", "yes", "same", "don't", "why", "sure", "much", "most", "i've", "didn't", "better", "being", "been", "around", "any", "another", "already", "against", "place", "whatever", "nowhere", "enough", "did", "use", "said", "should", "may", "which", "per", "get", "", "go", "could", "10", "might", "want", "used", "those", "good", "old", "new", "make", "things", "such", "it's", "you'll", "weren't", "haven't", "also", "won't","let's", "didn't", "yup", "mr"].

13.7 Appendix

Equivalent Terms were organized as a list of lists; each list element was a list of terms (at the 0 index) followed by one or more terms considered equivalent to (and programmatically replaced by) the 0 index term. Duplicate equivalent terms (like "president president") prevents a term like "president nixon" from being counted as 2 occurrences of the word "president" ("president president" is replaced as simply "president" in the text).

termList = [["heating oil", "heating fuel", "heat oil", "propane", "methane",
"natural gas"", "airplane fuel", "aviation fuel"],

["gasoline", "fuel", "petroleum", "ethel", "gasahol", "gas "],
["electricity", "electric power", "electrical", "electric"],
["alternative", "solar-powered", "wind-powered", "sun-powered",
"hydrogen", "solar"],
["alcohol", "ethanol"],
["dst", "daylight savings", "daylight saving", "daylight time",
"daylight saving time", "daylight savings time", "in the dark",
"day light savings", "day light saving", "day light", "it's dark"],
["conservation", "saved", "savings", "saving", "reduced", "reduction",
"cut down", "turn down", "conserved", "conserves", "conserve", "turn off",
"saves", "reduce", "save"],
["sanitizer", "hand sanitizer", "hand sanitizers", "disinfectant",
"disinfectants", "clorox"],
["president", "nixon's", "trump's", "richard r. nixon", "donald j. trump",
"richard nixon", "donald trump",
"the administration", "the administration's", "potus",
"nixon", "trump", "president president"],
["mph", "miles per hour", "miles an hour", "mile an hour",
"mile per hour", "m.p.h"],
["usa citizen", "americans", "american"],
["omaha", "city"],
["usa", "united states", "u.s.a.", "u.s.", "america"],
["nebraska", "state legislature", "state", "unicameral"],
["government", "washington", "lincoln"],
["arab", "arabs", "saudi arabia"],
["oil industry", "oil companies", "oil barons", "oil sheiks",
"oil company", "oil baron", "oil sheik"],
["panic", "hysteria", "crisis", "ordeal", "emergency"],
["microbes", "viruses", "bacteria", "germs", "a virus"],
["h1n1", "h1n1 virus","h1n1 epidemic", "h1n1 pandemic"],
["influenza", "flu virus", "flu epidemic", "flu pandemic", " flu"],
["covid", "coronavirus", "virus", "pandemic", "covid-19",
"epidemic", "covid's", "covid covid"],
["liberal", "liberals", "the left"],
["oppd", "omaha public power district", "electric company"],
["mud", "metropolitan utilities district", "gas company"],
["conservative", "conservatives", "the right"],
["congress", "senators", "representatives", "congressmen"],
["physician", "doctors", "physicians", "mds", "doctor"],
["elderly", "seniors", "senior citizen", "nursing home resident"],
["school", "schools", "classrooms", "classroom", "classes",
"schoolroom", "class"],
["worker", "employees", "workers", "professionals", "employee",
"professional"],
["mask", "masks", "face covering", "face coverings"],

[“mayor”, “stothert”, “zorinsky”, “mayors”, “mayor mayor”],
[“online”, “virtual”, “remote”, “on the computer”],
[“work”, “job”, “employment”],
[“media”, “newspaper”, “internet”, “world-herald”,
“television”, “tv”, “facebook”, “twitter”, “tweet”, “press”,
“publish”, “paper”, “news”],
[“editorial”, “comments”, “commented”, “comment”, “opinion”,
“opinions”, “public pulse”, “letter to the editor”,
“pulse”, “letter”, “letters”, “editorial editorial”],
[“governor”, “gov.”, “j. j. exon”, “ricketts”, “exon”,
“governors”, “governor governor”],
[“mandate”, “requirements”, “requirement”, “laws”, “law”,
“mandates”, “limits”, “limit”, “sanctions”, “sanction”],
[“guideline”, “guidelines”, “suggestions”, “suggestion”, “recommendations”,
“recommendation”, “requests”, “requested”, “request”],
[“hospital”, “medical”, “clinic”, “hospitals”, “clinics”, “medical facility”],
[“retailer”, “retailers”, “wholesaler”, “wholesalers”, “superstore”,
“superstores”, “malls”, “mall”, “shopping centers”, “shopping center”,
“supermarkets”, “supermarket”, “big box stores”,
“big box store”, “stores“, “store”, “retailer retailer”]].

References

Alexopoulos M, Cohen J (2015) The power of print: uncertainty shocks, markets, and the economy. Int Rev Econ Financ 40:8–28

Barsky RB, Basu S, Lee K (2015) Whither news shocks? NBER Macroecon Annu 29(1):225–264

Barsky RB, Sims ER (2011) News shocks and business cycles. J Monet Econ 58(3):273–289

Chen SH (2012) Varieties of agents in agent-based computational economics: a historical and an interdisciplinary perspective. J Econ Dyn Control 36(1):1–25

Dyas-Correia S, Alexopoulos M (2014) Text and data mining: searching for buried treasures. Ser Rev 40(3):210–216

Gentzkow M, Kelly B, Taddy M (2019) Text as data. J Econ Lit 57(3):535–574

Holmes DR (2019) Markets are a function of language: notes on a narrative economics (No. 2019–18). Economics Discussion Papers

Hsu C, Yu T, Chen SH (2021) Narrative economics using textual analysis of newspaper data: new insights into the US Silver Purchase Act and Chinese price level in 1928–1936. J Comput Soc Sci 1–25

Koslyk CAR (2019) News-driven business cycles: a narrative approach. Master’s Thesis, University of Tartu

Omaha World-Herald (1973) Immediate cut in Arab oil has little effect on U.S., p 2. (18 Oct 1973)

Omaha World-Herald (2020) Ricketts likens virus preparation to stocking up for a snowstorm, pp 1, 3 subtitle: virus: cancellation of large-scale events possible. (3 Mar 2020)

OWH (2020) Omaha World-Herald. https://infoweb.newsbank.com/

Paugam L, Stolowy H, Gendron Y (2020) Deploying narrative economics to understand financial market dynamics: an analysis of activist short sellers’ rhetoric. Contemp Account Res

Shiller RJ (2017) Narrative economics. Am Econ Rev 107(4):967–1004

Shiller RJ (2020) Narrative economics: how stories go viral and drive major economic events. Princeton University Press

Soroka SN, Stecula DA, Wlezien C (2014) It's (change in) the (future) economy, stupid: economic indicators, the media, and public opinion. Am J Polit Sci 59(2):457–474. https://onlinelibrary-wiley-com.cuhsl.creighton.edu/doi/epdf. https://doi.org/10.1111/ajps.12145

Volchik V (2017) Narrative and institutional economics. Available at SSRN 3621881

Chapter 14
Design and Performance Metrics for Autonomous Human-Machine Teams and Systems (A-HMT-S)

W. F. Lawless

Abstract Metrics for the design and operations of a team or organization are likely to use a rational approach organized around Shannon's theory of information. When these metrics are applied, uncertainty is commonly located in methods, algorithms, models or incomplete data. But reality is noisy, incomplete and uncertain. To function in reality, metrics for the design and performance of a team or an organization as presently constituted have to be transformed to assist autonomous systems with making decisions in uncertain environments, especially when the uncertainty is posed by opposing autonomous systems (e.g., conflict, deception, competition). We address this more difficult problem in our research with our theory of the interdependence of complementarity. Consequently, we discuss metrics derived from our prior findings that support separating team structure and performance, and metrics based on our new discovery that vulnerability is a target in an opponent's team or a key concern in one's own team or organization.

Keywords Interdependence · Complementarity · Bistability · Uncertainty and incompleteness · Non-factorable information and tradeoffs · Metrics for design and performance · Geometry · Vulnerability

14.1 Introduction

The pursuit of computational autonomy for human–machine teams and systems, while difficult, is showing the first signs of success across the economy and military. For example, in the event of an emergency occurring during a flight, an "auto-land" technology for single-piloted aircraft has been developed that may have application within a decade to reshape commercial jets, military fighter jets, and military helicopters (Pasztor 2020). There are also signs of more mundane applications; e.g., the use of artificial intelligence (machine learning) that allows credit or debit transactions to continue autonomously in the event that a bank's network has been shut

W. F. Lawless (✉)
Professor of Math and Psychology, Paine College, Augusta, GA 30901, USA
e-mail: w.lawless@icloud.com

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_14

down, limited to preventing what amounts to an inconvenience (Beatrice 2020). But to develop autonomy for human–machine teams and systems addressing complex problems in reality requires an approach to interdependence that has yet to be developed computationally.

For our theory of the interdependence of complementarity applied to autonomous human–machine teams and systems (A-HMT-S), we briefly review our research for metrics covering interdependence (bistability, incompleteness, and non-factorability); degrees of freedom; Shannon's information theory and independence; von Neumann's entropy theory and non-interdependence; structural entropy production and performance entropy production; the whole and its parts; and we extend our past research to the recent discovery of organizational vulnerability uncovered during competition.

14.2 Justification for Interdependence Theory

Rational models of human autonomy aim to predict, possibly control, human behavior. There are two primary models, the cognitive model that treats behavior as implicit, and the behavioral model that treats beliefs as implicit. The cognitive model reigned until validity issues arose, including Axelrod's prediction that cooperation produces the best outcomes for societies (Rand and Novak 2013; compare p. 413 and p. 422). And while dismissing the value of beliefs improved predictions of behavior, improvement occurred in situations where beliefs were suppressed, unimportant, or when enforced cooperation accompanied low risk, certain or isolated environments. For our purposes, rational models lack the supporting evidence to make successful predictions in the wild (e.g., the failure of super forecasters to predict Brexit or Trump[1]; cf. Tetlock and Gardner 2015), impeding generalizations to artificial intelligence (AI). Nor can these models scale to teams or systems, another flaw.

However, rational models fail in the presence of conflict and environmental or situational uncertainty, their fatal flaw, precluding a metrics of performance based on rational models. These shortcomings leave rational models ill-prepared to assist the technical revolution proposed by autonomous Human–Machine Teams or systems (A-HMT-S).

To integrate traditional models of physical behavior and those of cognition for the development of A-HMT-S, we have developed the interdependence theory of complementarity, often overlooked because of the bewilderment interdependence causes in the laboratory (specifically, "bewildering complexities"; in Jones 1998, p. 33; however, Cooke and Hilton 2015, reviewed studies which have revitalized the study

---

[1] Tetlock's website: http://goodjudgment.com/superforecasting/index.php/2016/11/03/is-donald-trump-mr-brexit/.

of interdependence). Where the rational model fails in the face of conflict or uncertainty, complementarity has thrived. The best human science teams are fully interdependent; the interactions of teammates is a team's intelligence (specifically, not located in the intellects of independent members; in Cooke 2020); and complementarity is quantum-like; that is, per Bohr (1955), complementarity as a concept operates like Heisenberg's uncertainty principle to produce a "unity of knowledge" for predictions about human behavior, but without its mathematical exactness (Lawless 2020). For example, with our model, we have reported that, facing external uncertainty or conflict, human debate exploits the interdependent bistable views of reality with tradeoffs seeking the best path forward. Explaining uncertain contexts, which no single agent or team can determine alone (Lawless et al. 2019), necessitates that the members of an A-HMT-S express their actions in causal terms however imperfectly. Our approach has led to two discoveries, which both generalize and scale:

1. Structure and performance entropy production are complementary; and,
2. The vulnerabilities generated during competition provide effective metrics for the design and performance of a team or an organization for social evolution, invisible to the theories and practices of cooperation.

We consider these issues and also note that the metrics are a natural basis for the geometric study of information theoretic issues. At this point, we are still considering the various geometric structures that best capture these issues. We note that Fisher information has been used by us and many others for discussions on these matters (Moskowitz et al. 2020).

14.3 A New Approach With Metrics

Traditionally (Conant 1976), when a group is composed of independent individuals, the whole is aggregated as the sum of its individual contributions, represented by Shannon information and Eq. (14.1), where S is used for entropy, and with A, B, and A, B for agent A, agent B and team AB:

$$S_{A,B} = S_A + S_B \tag{14.1}$$

Equation (14.1) can be revised to Eq. (14.2) for the case when the individuals are not completely independent, giving:

$$S_{A,B} \geq S_A, S_B \tag{14.2}$$

Equation (14.2) tells us that the joint entropy for a bipartite system is at least equal to one of its members. But there is more to be learned with Eq. (14.3):

$$S_{A|B} = S_{A,B} - S_B \geq 0 \tag{14.3}$$

Equation (14.3) addresses any lingering ignorance about *A* that we may have when we know *B*, and equals zero when knowing *B* makes us certain about *A* (Preskill 2016).

Conant's (1976) conclusion with Shannon's information theory is that interdependence should be minimized. From Social Psychology, Kenny et al. (1998, p. 235) offer a statistical correction to make interdependent data independent ostensibly for the purpose of improving the replicability of an experiment. But with the replication and validation of many social science concepts either overthrown or in question (Nosek 2015), we have argued that Kenny's correction to remove any interdependence that is found in the data is akin to treating quantum effects at the atomic level as an annoyance rather than the resource that they have become.

Contradicting Conant and Kenny, the recent literature suggests that the best teams are highly interdependent (Cummings 2015; Cook and Hilton 2015). If true, introducing interdependence into human machine teams is a critical step on the path to solving, and providing metrics for, the computational autonomy of A-HMT-S that are planned to be operated in the wild (Lawless et al. 2019).

In prior research (Lawless 2019b), we have arbitrarily divided interdependence into bistability (e.g., two sides to every story; two tribes with incommensurable beliefs; and an agent can either act as an individual or as a member of a group, but not both simultaneously); incompleteness (i.e., believing there is only one side to a story necessarily diminishes the value of an opposing side, commonly associated with mistakes, accidents or tragedies; in Sofge et al. 2019); and non-factorability (e.g., under uncertainty or facing conflict, a business tradeoff; a competing alternative decision; also, most social situations are difficult to unravel objectively, often requiring, for example, a court of inquiry to disentangle a severe social dispute).

14.4 Field Results

We have found that redundancy impedes the performance of a team for the top oil firms in the world (Lawless 2017a), replicated for the top militaries around the globe (Lawless 2017b), leading us to conclude that redundancy occurs more under authoritarian than democratic governments, suggesting that redundant workers are more likely associated with corruption and not with effectiveness nor efficiency, even when faced by the shock to the stock-market caused by the pandemic in 2020 (Lawless 2020). We have generalized the results for autonomous human–machine teams (A-HMTs; in Lawless et al. 2019), for which we have made a bet on the predicted usurpation of autonomy by a machine when human members in the team have become inept, malfeasant, or act with a criminal intent to harm other humans (Sofge et al. 2019).

Subsequently, we considered the value of an education for a nation's production of patents. For the study (Lawless 2019a, b), we used the public data from the United Nations' Human Development Index for all 19 Middle-Eastern North African (MENA) nations plus Israel and found a significant result, providing support for well-educated research scientists. In contrast, however, in other research for the US Air Force, we found no association between a fighter pilot's education for air-combat maneuvering and performance in air-to-air combat (reviewed in Lawless 2017a,b). These two findings illustrate the orthogonality in comparing patent productivity, with its emphasis on the best intellectual education, versus air combat, with its emphasis on extreme physical training. Both reflect the top performance characterized by maximum entropy production (MEP; in Martyushev 2013), but this maximum productivity is orthogonal across the two fields of endeavor.

For our next project, we constructed a social harmonic oscillator as a model of how free humans spontaneously reach decisions by processing information circulating driven by debate or other form of competition (e.g., Lawless 2020); in this case, the U.S. Department of Energy was planning to close two highly radioactive waste (HLW) tanks, but had to obtain agreement with the Nuclear Regulatory Commission (NRC), which had repeatedly asked DOE to redesign its closure process in an oscillation that continued for seven years. Frustrated, the State of South Carolina complained to DOE's citizens advisory board (CAB) at DOE's Savannah River Site that DOE was going to miss its legally mandated milestone to close these two HLW tanks in the State. The SRS-CAB called the two agencies together and demanded the closure of these two HLW tanks; both agencies agreed in what became not only their fastest joint decision ever made on HLW tanks, but also an agreement that led to their immediate closure (Lawless et al. 2014).

An interdependent team or organization is commonly composed of orthogonal roles, where each member of a team is dependent on the others in the team to achieve individual and team success (e.g., a minimally sized restaurant has a cook, a waiter and a cashier in orthogonal roles). In the limit, we hypothesize that a team composed of orthogonal elements (agents) can become a structural unit. When successful and when it functions as a single unit, its structural entropy production becomes minimal (least entropy production or LEP; Lawless 2020). As a Boltzman-like reduction in the degrees of freedom occurs (i.e., an individual system with *N* degrees of freedom becomes a macro-team of one with *1/N* degrees), we can represent this state for a structure's degrees of freedom (*dof*) in the limit with Eq. (14.4):

$$S_{A,B} = lim_{dof \to 1} log(dof) = 0 \tag{14.4}$$

When LEP occurs for a team's structure, counterintuitively, the "whole is greater than the sum of its parts." This well-known state was first recognized in quantum mechanics and named "entanglement" by Schrödinger (1935) to represent a loss of independence between the parts of a whole, but it has also been reported in social psychology (Lewin 1951) and Systems Engineering (Walden et al. 2015). Specifically, from Schrödinger (1935, p. 555),

> the best possible knowledge of a *whole* does not necessarily include the best possible knowledge of all its *parts,* even though they may be entirely separate and therefore virtually capable of being 'best possibly known,' i.e., of possessing, each of them, a representative of its own.

> The lack of knowledge is by no means due to the interaction being insufficiently known—at least not in the way that it could possibly be known more completely—it is due to the interaction itself.

From the Systems Engineering Handbook (Walden et al. 2015), "A System is a set of elements in interaction" (Bertalanffy 1968) where systems "… often exhibit emergence, behavior which is meaningful only when attributed to the whole, not to its parts" (Checkland 1999).

But Schrödinger (1935, p. 555) added that:

> Attention has recently been called to the obvious but very disconcerting fact that even though we restrict the disentangling measurements to *one* system, the representative obtained for the *other* system is by no means independent of the particular choice of observations which we select for that purpose and which by the way are *entirely* arbitrary.

To reflect this state of the "whole," when the lack of independence among its parts occurs, we can use von Neumann's Eq. (14.5) which is sub-additive because independence can no longer be assumed:

$$S_{A,B} \leq S_A + S_B \tag{14.5}$$

Equation (14.5) accounts for the many problems with implicit behavior in the cognitive model that impedes validation (see Thagard 2019) and implicit beliefs in the behavioral model that impede scaling (e.g., Liu and Barabasi 2016). In both of these models, the little recognized problem is represented with Eq. (14.5) as the loss of information about a "whole" team or system about their parts, contradicting the basic assumptions of the rational model and independence in the cognitive model.

Equations (14.4) and (14.5) encouraged us to separate a team's task performance from the team's structure in agreement with Salas' research (in Bisbey et al. 2019). This separation allowed us to recognize that the best performing teams are devoting minimal free energy to its team structure and most of its free energy to operate at maximum entropy production (MEP; from Martyushev 2013). As a relevant example, we can characterize MEP by searching fully across a solution space to perform at the highest level; viz., our earlier discussion about producing a patent (Lawless 2019a). We have assumed that teams operating at MEP require a team functioning with the highest level of team intelligence (Wissner-Gross and Freer 2013), where intelligence is exhibited through the interactions of a team (Cooke 2020; Cooke and Hilton 2015).

Complementarity offers a natural tradeoff between two conjugate factors, which we have assumed to be LEP and MEP, respectively, at a minimum for the structure and a maximum for the task, both contingent on the free energy available (Lawless 2020). To achieve MEP requires a stable structure (Martyushev 2013). But teams often face other teams in states of competition. This realization led us to the discovery that in a state of competition, when a team discovers a vulnerability in its opponent, it can exploit that vulnerability to defeat its opponent (Lawless 2020). We conceptualize the information produced during competition as flowing from *Team C* to *Team D* and back again in a state of oscillation (e.g., as in a debate; this oscillation may not have limit points). In response to processing this incoming information from its opponent's actions, it is this circulating information that a team uses for its decision on how, when and where to attack its opponent. If *DA* represents a decision advantage, and τ represents the power of a team's decision during an oscillation, then

$$DA_{Team-C-over-Team-D} = \tau_{output}/\tau_{input} \quad (14.6)$$

To become or remain competitive, if one team makes its decisions on its incoming information faster than its opponent, ceteris paribus, Eq. (14.6) has informed us that the essence of speed in making a decision is in exploiting the vulnerability in an opponent's or in one's own organization, motivating the need to replace unnecessary (redundant) or weaker teammates with the best skilled participants possible as the members of a team, the mergers for an organization, or the adoption of new technology.

14.5 Vulnerability

If an attack is effective against a team at its structure's point of vulnerability, its structure's increase in entropy production will be taken from the team's supply of free energy, reducing its MEP performance. The war of words between Apple and Facebook provides an excellent example of targeting a structural vulnerability in Facebook used by Apple and publicized during the aggressive competition between these two firms. As reported in the *Wall Street Journal* (Haggin et al. 2021),

> Facebook Inc. will suffer damage to its core business when Apple Inc. implements new privacy changes, advertising industry experts say, as it becomes harder for the social-media company to gather user data and prove that ads on its platform work. The core of Facebook's business, its flagship app and Instagram, would be under pressure, too. The Apple change will require mobile apps to seek users' permission before tracking their activity, restricting the flow of data Facebook gets from apps to help build profiles of its users. Those profiles allow Facebook's advertisers to target their ads efficiently. The change will also make it harder for advertisers to measure the return they get for the ads they run on Facebook—how many people see those ads on mobile phones and take actions such as installing an app, for example.

Two organizations often struggle to find an advantage or a vulnerability in each other that can be made into a competitive advantage. For example, from *CNBC Markets* (Sheetz 2021),

> The satellite internet projects of the two richest men on the planet continue to spar behind the scenes with federal regulators, with Amazon on Thursday clarifying its position in response to recent accusations from Elon Musk and SpaceX that Jeff Bezos' company is attempting to "stifle competition" in the sector. Representatives of Amazon spoke to Federal Communications Commission officials earlier this week, doubling down on its position that the FCC should not approve SpaceX's modification request for parts of its Starlink satellite network. Amazon and SpaceX are working to build space-based internet networks—called Kuiper and

> Starlink, respectively—by launching thousands of satellites into orbit, known in the industry as a constellation.

The opposite happens when a company is aware of its own vulnerability, such as when it carries an unsustainable debt load, making it unprepared for an unexpected and subsequent national or global calamity, like the Covid-19 pandemic which has caused several airlines to experience severe problems. This collapse happened recently with China's HNA Airlines. From the *Financial Times* (Hale 2021),

> Creditors of struggling Chinese conglomerate HNA Group have applied for bankruptcy proceedings after a court said the company was unable to pay its debts. ... The company, which started life as an airline and took on huge debt to expand into one of China's flagship global conglomerates, said it would co-operate and "support the court to protect the legal rights and interests of creditors." Over recent years, HNA has struggled with increasingly unmanageable debt levels following a series of high-profile overseas investments, including stakes in Deutsche Bank and the Hilton Worldwide hotel chain.

An example of a vulnerability that the company could not solve is apparent when a company spins off or divests one of its units or assets. Divestiture is happening with the Kraft Heinz company after it struggled with shifting preferences among consumers. *The Street* (Owusu 2021) recently reported that Kraft Heinz is planning to "sell its Planters business to Hormel Foods … [after the firm has] struggled since Kraft and Heinz merged five years ago, with analysts seeing it as too slow to alter older brands and unwilling to commit adequate resources to marketing."

Another example concerns the rush into a new market by many firms at a time when a market cannot support all of the competitors, creating a glut that sets the stage for a consolidation in the market. An excellent example is the small-rocket market. From the *Wall Street Journal* (Pasztor 2021),

> Companies and entrepreneurs world-wide are working on more than 100 new small-rocket ventures, but industry officials anticipate a shakeout eventually may leave just a handful of survivors. ... All [of these] competitors ... confront the same market reality. Burgeoning corporate, civilian and military uses of compact satellites weighing under a ton—ranging from toaster-size models to versions resembling refrigerators—won't generate enough demand to support the current glut of small launchers.

A last one of many examples available concerns the motivation to merge with a small firm to obtain access to its technology to keep pace in an arms race against major competitors. Patel and Bruell (2021) report that Walmart is chasing large competitors Amazon, Facebook and Alphabet's Google, three firms with a large head start:

> Walmart Inc. will acquire technology from Thunder Industries, a company that uses automation to create digital ads, as it continues to invest in its ad business and seeks a greater slice of marketing budgets from small businesses. Walmart instead will use Thunder's technology to launch a self-service tool that helps advertisers make and buy numerous versions of display ads targeting different kinds of consumers on its properties. Walmart ... [is] building digital-ad offerings to generate new revenue by using their shopper data to help marketers target customers online and in stores.

14.6 Vulnerability from the use of Deception

Ancient Greece. Vulnerability can be created in an opponent with the use of deception. One of the first uses recorded was by Themistocles (Friedman 2015), leader of the Athenian fleet during the Peloponnesian War and in the Battle of Salamis in 480 BC. Pretending to side with the Persians, Themistocles sent a message to Xerxes, leader of the Persians, that Xerxes should send his fleet to blockade the opening of the Bay of Salamis, trapping the Greek fleet. Themistocles had promised to defect from the Greeks once the battle began and to attack the other ships in the Athenian fleet. According to Herodotus, however, the narrow confines of the bay would nullify the numerical advantage of the Persian fleet, allowing the superior Greek seamanship and ship boarding to defeat Xerxes. Once the Persian fleet entered the restrictive waters of the Bay, the Greek fleet retreated to further encourage the Persian Navy. The Persian naval fleet was destroyed and routed. Once the Persians lost cohesion and began to flee, the Persian disaster unfolded as Themistocles had planned.

China. General M. Hayden, a former Central Intelligence Agency (CIA) and National Security Administration (NSA) chief, stated that the Chinese stole millions of records from federal employees for the innovativeness that had eluded China. He told his Chinese counterparts: "You can't get your game to the next level by just stealing our stuff. You're going to have to innovate" (Baker 2015).

China, however, has neutralized some of its economic advantage and its ability to innovate. Consider that the R&D expenditures by China are second in the world to the U.S. But China's state directed finance, its weak intellectual property protections and its rampant corruption impede innovation. In China (Taplin 2018), "Small private-sector firms often only have access to capital through expensive shadow banking channels, and risk that some better connected, state backed firm will make off with their designs–with little recourse."

The former Director of National Intelligence, Ratcliffe (2020), wrote that the Chinese Communist Party engages in economic espionage to "rob, replicate and replace" companies of their intellectual property, replicate the technology, and then replace the U.S. firms in the global marketplace. Take Sinovel, for example:

> In 2018 a federal jury found the Chinese wind-turbine manufacturer guilty of stealing trade secrets from American Superconductor. Penalties were imposed but the damage was done. The theft resulted in the U.S. company losing more than $1 billion in shareholder value and cutting 700 jobs. Today Sinovel sells wind turbines world-wide as if it built a legitimate business through ingenuity and hard work rather than theft.

By applying our findings to China, its social systems replete with redundant individuals will be relatively full of vulnerabilities, leading to an ever-tightening round of social-political-military control, reflected by China's use of re-education work camps (e.g., ED 2020). But there is also positive signs that China is its regulators are moving to rectify "one of the world's worst polluters" (Hua 2021):

> The environmental report was unusually harsh: China's National Energy Administration had skirted air-pollution regulations, allowed excess construction of coal plants and failed to

> heed leader Xi Jinping's instructions on environmental matters. ... The rare public rebuke—authorized by two of China's top state authorities, the Chinese Communist Party's Central Committee and the country's cabinet—signaled to some that environmental regulators in one of the world's worst polluters had been empowered to counter the country's entrenched coal industry.

14.7 Future Research

We arrive at our proposed metrics for the design and performance of a team or an organization for the complementarity between two conjugate-like variables[2] to relate uncertainty in the free energy available directed at the forces binding a team's structure and driving its MEP.

We now consider a team to be a combined mechanical-thermodynamic system, with an ability to think, modeled somewhat like Maxwell's Demon, where the team's thinking is assumed to focus on the tradeoff between the conjugate variables involved in directing its free energy to keeping the team's structure stable or to maximizing its performance (the tradeoff is manifold and complex; it is internal for a team and external for the team's opponent). In this case, the product of the two quantities that are conjugate-like have units of free energy. Let us further imagine that, for Eq. (14.7) below, which we treat as a work-in-progress, we have a single uncertainty relationship of conjugate-like variables with the force of keeping structure stable *times* (in exchange for the maximum flow of) the entropy flowing through the system that represents its productivity, the product being greater than or equal to the free energy available to a system. Let us also assume that, for our system, a small increment of energy is the product of a force, F, times a small displacement, the greater the displacement, the greater the force. In our desired system, as long as the displacement is minimal or zero, most of the team's free energy is available to produce a flow of entropy, P, that maximizes MEP. In turn, if all of the available free energy is directed at maximizing MEP, where a displacement from MEP produces a force, assuming that it is at a maximum, we speculate that the product produces uncertainty in the free energy, A, such that:

$$\begin{aligned} \Delta F_{Force-structure-entropy} * \Delta P_{MEP-performance-entropy} \\ \geq C_{Unknown} =? = \Delta A_{free-energy} \end{aligned} \tag{14.7}$$

[2] Conjugate variables are pairs of variables defined mathematically to become Fourier transform duals that lead naturally to uncertainty relations. In contrast, conjugate-like variables form the basis of Bohr's (1955) concept of complementarity (Lawless 2020).

14.8 Conclusion

Based on the foregoing relationships and evidence from the field (of which is too numerous to discuss at length in this chapter), instead of supporting the anti-Darwinism of Hare and Woods (2020), and contradicting the often overstated value of social cooperation (e.g., Johnson and Johnson 2015; Axelrod 1980, pp. 7–8), our theory of the interdependence of complementarity offers support for social evolution (e.g., political, commercial, military). Moreover, from McFadden and Al-Khalili's (2018; see also Cao et al. 2020) review of quantum biology:

> Recent work has demonstrated that the retention of quantum dynamics in biological systems is intricately connected with environmental fluctuations taking place at biologically relevant length and time scales.

If we apply McFadden and Al-Khalili directly to competition with our theory of the complementarity of interdependence, it indicates to us that complementarity in teams is more likely to occur under the noise and strife of competition that makes a team work harder to find the information available to keen observers needed to guide their own competitive (re)actions, rather than the cooperation which consumes or hides information from observers. In this evolving, complex environment, metrics for the design and performance of a team or organization have a role to play.

When the situation reflected by Eq. (14.5) occurs, there is full information available about the whole team's productivity, but a loss of information about how its parts are interacting, making the whole's production greater than the sum of its parts. However, based on Preskill's (2016) argument, instead, we argue that the lost information is located in the correlations among the interactions of the members of a team. For support, we look to Cooke (2020) and her team who reported that the intelligence of a team that improves its performance is not found in the knowledge possessed by a team, but in their *interactions*, which are interdependent.

Quantum theory and interdependence theory both model state-dependent phenomena, but quantum theory is governed by a precise mathematics, while interdependence theory is governed by concepts, similar to Bohr's (1955) theory of complementarity. Despite this reservation, the findings from interdependence theory serve to promote its application to autonomous human–machine teams and systems. Its mathematics is not as precise as matching learning, nor quantum mechanics, but neither of those sciences have been able to demonstrate their successful application to autonomy. With interdependence theory, in autonomous systems, we have predicted and found that human firms resilient to shocks have minimum redundancy (Lawless 2020).

Acknowledgments I would like to thank Ira S. Moskowitz, Mathematician, Information Technology Division, the Naval Research Laboratory, Washington, DC, for his many helpful suggestions.

References

Axelrod R (1984) The evolution of cooperation. Basic, New York

Baker G (2015) Michael Hayden Says U.S. Is Easy Prey for Hackers. Former CIA and NSA chief says 'shame on us' for not protecting critical information better, interview by the Wall Street Journal's editor in chief (2015, 6/21). http://www.wsj.com/articles/michael-hayden-says-u-s-is-easy-prey-for-hackers-1434924058

Beatrice A (2020) "Visa Kicks off AI feature that approves and declines card transactions. Analytics Insight, retrieved 12/22/2020 (2020, 9/23). https://www.analyticsinsight.net/visa-kicks-off-ai-feature-approves-declines-card-transactions/

von Bertalanffy L (1968) General system theory: foundations, development, applications, (rev. ed.) Braziller, New York

Bisbey TM, Reyes DL, Traylor AM, Salas E (2019) Teams of psychologists helping teams: the evolution of the science of team training. Am Psychol 74(3):278–289

Bohr N (1955) Science and the unity of knowledge. In: Leary L (ed) The unity of knowledge. Doubleday, New York, pp 44–62

Cao et al (2020) Quantum biology revisited, Science Advances 6(14):eaaz4888. https://doi.org/10.1126/sciadv.aaz4888

Checkland P (1999) Systems thinking, systems practice. Wiley, New York

Conant RC (1976) Laws of information which govern systems. IEEE Trans Syst Man Cybern 6:240–255

Cooke N (2020) Effective human-artificial intelligence teaming, AAAI-2020 Spring Symposium, Stanford, CA. (2020, 3/24). (cf. https://aaai.org/Symposia/Spring/sss20symposia.php#ss03)

Cooke NJ, Hilton ML (eds) (2015) Enhancing the effectiveness of team science. Authors: committee on the science of team science; national research council. National Academies Press, Washington (DC)

Cummings J (2015) Team science successes and challenges. National science foundation sponsored workshop on fundamentals of team science and the science of team science (June 2), Bethesda MD

ED (2020) The lesson in China's cotton slave trade, Washington Examiner. https://www.washingtonexaminer.com/opinion/editorials/the-lesson-in-chinas-cotton-slave-trade. Accessed 20 Dec 2020. (2020, 12/20)

Friedman BA (2015) The battle of salamis: themistocles and the birth of StrBridge. https://thestrategybridge.org/the-bridge/2015/9/29/the-battle-of-salamis-themistocles-and-the-birth-of-strategy. Accessed 29 Dec 2020. (2015, 9/29)

Haggin P, Hagey K, Schechner S (2021) Apple's privacy change will hit Facebook's core Ad business,. Wall Street J. https://www.wsj.com/articles/apples-privacy-change-will-hit-facebooks-core-ad-business-heres-how-11611938750. Accessed 30 Jan 2021

Hale T (2021) Creditors of Chinese conglomerate HNA begin bankruptcy proceedings. Company has struggled with increasingly unmanageable debt levels following high-profile investments. Financial Times. https://www.ft.com/content/0bd72906-e3eb-444d-a62b-6e16daf49ed0. Accessed 30 Jan 2021. (2021, 1/29)

Hare B, Woods V (2020) Survival of the friendliest. Understanding our origins and rediscovering our common humanity. Penguin/Random House

Hua S (2021) China hints at a shift on climate. A rare public rebuke of the energy agency signals to some that environmental regulators have been empowered to counter the coal industry. Wall Street J. https://www.wsj.com/articles/china-hints-at-a-shift-on-climate-11613125804. Accessed 14 Feb 2021. (2021, 2/12)

Johnson DW, Johnson RT (2015) Cooperation and competition. In: International encyclopedia of the social & behavioral sciences, 2nd edn, pp 856–861. https://doi.org/10.1016/B978-0-08-097086-8.24051-8

Jones EE (1998) Major developments in five decades of social psychology. In: Gilbert DT, Fiske ST, Lindzey G (eds) The handbook of social psychology, vol I. McGraw-Hill, Boston, pp 3–57

Kenny DA, Kashy DA, Bolger N (1998) Data analyses in social psychology. In: Gilbert DT, Fiske ST, Lindzey G (eds) Handbook of social psychology, 4th edn., vol 1. Boston, MA, McGraw-Hill, pp 233–265

Lawless WF, Akiyoshi M, Angjellari-Dajcic F, Whitton J (2014) Public consent for the geologic disposal of highly radioactive wastes and spent nuclear fuel. Int J Environ Stud 71(1):41–62

Lawless WF (2017a) The entangled nature of interdependence Bistability, irreproducibility and uncertainty. J Math Psychol 78:51–64

Lawless WF (2017b) The physics of teams: Interdependence, measurable entropy and computational emotion. Front Phys 5:30. https://doi.org/10.3389/fphy.2017.00030

Lawless WF (2019a) Interdependence for human-machine teams. Foundations of science, Springer. https://doi.org/10.1007/s10699-019-09632-5

Lawless WF (2019b) The interdependence of autonomous human-machine teams: the entropy of teams, but not individuals, advances science. Entropy 21(12):1195. https://doi.org/10.3390/e21121195

Lawless WF (2020) Quantum-like interdependence theory advances autonomous human–machine teams A-HMTs. Entropy, 22(11):1227. https://doi.org/10.3390/e22111227

Lawless WF, Mittu R, Sofge DA, Hiatt L (2019) Editorial, "Artificial intelligence (AI), autonomy and human-machine teams: Interdependence, context and explainable AI." AI Mag 40(3):5–13

Lewin K (1951) Field theory of social science. Selected theoretical papers. In: Cartwright D (ed). Harper & Brothers, New York

Liu YY, Barabási AL (2016) Control principles of complex systems. Rev Mod Phys 88(3):1–58. https://doi.org/10.1103/RevModPhys.88.035006

Martyushev LM (2013) Entropy and entropy production: old misconceptions and new breakthroughs. Entropy 15:1152–1170

McFadden J, Al-Khalili J (2018) The origins of quantum biology. In: Proceedings royal society A, vol 474, pp 20180674. https://doi.org/10.1098/rspa.2018.0674, rspa.royalsocietypublishing.org. Accessed 14 Jan 2020

Moskowitz IS, Russell S, Lawless WF (2020) An information geometric look at the valuing of information. In: Lawless, WF, Mittu R, Sofge DA (eds) Human-machine shared contexts. Elsevier

Nosek B, corresponding author from OCS (2015) Open collaboration of science: estimating the reproducibility of psychological science. Science 349(6251):943; supplementary: 4716-1–4716-9. The National Academies Press, Washington, DC. https://doi.org/10.17226/25303.

Owusu T (2021) Kraft Heinz reportedly near $3B deal to sell planters to Hormel. Kraft Heinz as soon as next week might deal its Planters snack business to Hormel for about $3 billion, a media report says. The Street. https://www.thestreet.com/investing/kraftheinz-reportedly-nearing-sale-of-planters-to-hormel. Accessed 8 Feb 2021. (2021, 2/2)

Pasztor A (2020) Press a button and this plane lands itself. Garmin's 'auto-land' capabilities take over in the event of a pilot's medical emergency. The software could start reshaping jetliner cockpits within the decade, proponents say. Wall Street J. https://www.wsj.com/articles/press-a-button-and-this-plane-lands-itself-11595044800. Accessed 22 Dec 2020. (2020, 7/18)

Pasztor A (2021) Rough landing expected for the glut of new small-rocket makers. Increased use of compact satellites by companies and military isn't seen supporting the current number of smaller launchers. Wall Street J. https://www.wsj.com/articles/rough-landing-expected-for-the-glut-of-new-small-rocket-makers-11612710000. Accessed 8 Feb 2021. (2021, 2/7)

Patel S, Bruell A (2021, 2/4) Walmart buys Ad Tech to chase small-business advertisers. Major retailers continue efforts to become powers in ad sales. Wall Street J. https://www.wsj.com/articles/walmart-buys-ad-tech-to-chase-small-business-advertisers-11612438200. Accessed 8 Feb 2021. (2021, 2/4)

Preskill J (2016) Quantum information. In: Quantum Shannon theory, Ch. 12. California Institute of Technology. http://www.theory.caltech.edu/people/preskill/ph219/. Accessed 10 Feb 2021

Rand DG, Nowak MA (2013) Human cooperation. Cogn Sci 17(8):413–425

Ratcliffe J (2020) China is national security threat no. 1. Resisting Beijing's attempt to reshape and dominate the world is the challenge of our generation. Wall Street J. https://www.wsj.com/articles/china-is-national-security-threat-no-1-11607019599. Accessed 4 Dec 2020. (2020, 12/3)
Schrödinger E (1935) Discussion of probability relations between separated systems. In: Proceedings of the Cambridge philosophical society, vol 31, pp 555–563. (vol 32, pp 446–451, 1936)
Sheetz M (2021) Amazon responds to Elon Musk's accusations of impeding SpaceX's Starlink satellite internet plans. CNBC markets. https://www.cnbc.com/2021/02/05/amazons-kuiper-responds-to-elon-musks-spacex-on-fcc-request.html. Accessed 7 Feb 2021. (2021, 2/5)
Sofge DA, Lawless WF, Mittu R (2019) AI Bookie Bet: autonomous AI takes responsibility: threat, savior or both? AI Mag 40(3):79–84
Taplin N (2018) Can China's red capital really innovate? U.S. technology theft from Britain helped k2018, ick-start the industrial revolution on American shores. Will China be able to replicate that success? Wall Street J. (2018, 5/14). https://www.wsj.com/articles/can-chinas-red-capital-really-innovate-1526299173
Tetlock PE, Gardner D (2015) Superforecasting: the art and science of prediction, Crown
Thagard P (2019) Cognitive science. In: Zalta EN (ed) The Stanford encyclopedia of philosophy. https://plato.stanford.edu/archives/spr2019/entries/cognitive-science. Accessed 15 Nov 2020
Walden DD, Roedler GJ, Forsberg KJ, Hamelin RD, Shortell TM (eds) (2015) Systems engineering handbook, 4th edn. INCOSE-TP-2003-002-04. Wiley, Hoboken, NJ
Wissner-Gross AD, Freer CE (2013) Causal entropic forces. Phys Rev Lett 110(168702):1–5

Chapter 15
Thirty-Five Years of Computational Economics

Robert Marks

Abstract This chapter describes the evolution of my work in computational economics, from 1987 to 2020, as I submitted algorithms to play a generalisation of the iterated Prisoner's Dilemma (IPD), applied a new method of machine learning (the genetic algorithm, or GA) to this, extended these techniques to trying to understand and improve on asymmetrical seller behaviour in historical oligopoly pricing, generalised this to fully fledged agent-based models (ABMs), applied the GA to exploring the best method of decision making in uncertain situations, and finally as I used simulation to search exhaustively among decision-making methods with uncertainty. This research program also led me to derive new techniques for validating the output of simulation models when the metric was *ordinal* (rather than the *interval* or *ratio* metrics usually encountered, especially in physics and engineering). At least since 1995, while undertaking this work, I have been involved with Distinguished Professor Shu-Heng Chen at many conferences, and with several publications and lectures. This chapter recounts this history.

Keywords Decision-making · Strategic oligopoly · Simulation · Validation of patterns

I'd like to thank Chris Adam, Bob Wood, Peter McBurney, David Midgley, Nick Vriend, Robert Axelrod, Leigh Tesfatsion and Shu-Heng Chen for their comments and suggestions.

R. Marks (✉)
Economics, University of New South Wales, Sydney, Australia
e-mail: robert.marks@gmail.com

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_15

15.1 Tournaments to Explore the Iterated Prisoner's Dilemma

In 1987 I won the Second MIT Competitive Strategy Tournament.[1] This was a three-person pricing competition among imperfect substitutes, in which pricing low would increase sales but cut into profits. Just what my win led to, including my connection with Shu-Heng Chen, is recounted below.

My 1978 Ph.D. thesis had been a theoretical study of the interactions of markets for output, labour, and non-renewable resources in the macro economy.[2] This led to publications on energy and the environment. It did not lead directly to the work described below. From my teaching[3] and interest in environmental issues, I was, however, taken with the Prisoner's Dilemma (Marks 1998), thanks to Hardin's (1968) "The tragedy of the commons" and Schelling's (1978) *Micromotives and Macrobehavior*. As computers and computing had become more than a means of number crunching, and as the Internet had spread and enabled world-wide communication (even before the World Wide Web), the notion that it was possible to pit submitted algorithms to compete in a tournament, and that researchers from around the world could submit algorithms to compete was new.

Political scientist Robert Axelrod had been the first to use this possibility seriously, the first on-line instance of what we now call crowdsourcing. In 1979 he invited submission of algorithms via email to play an IPD game in silico, in order to see whether the temptation to defect (in the one-shot game) could be avoided in the iterated game.[4]

The Prisoner's Dilemma is an example of a strategic situation where, in the one-off interaction, mutual defection D, D is the likely outcome, whereas mutual cooperation C, C is Pareto superior (Marks 1998). That is, the payoff to each player from C, C is higher than that from D, D. Could repetition result in C, C? This was a contention of Axelrod's, who spelt this out in his 1984 book which asked how cooperation might have evolved in political systems.

His research had begun in the late 1970s, with an investigation of the emergence of cooperative behaviour and social norms in Hobbesian societies. In Axelrod's first tournament, Anatol Rapoport submitted a simple algorithm: start by cooperating and then mimic the other player's action in the previous round. This strategy, now known as Tit for Tat, was very successful, outperforming all other strategies but for Always Defect (Axelrod 1980a).

---

[1] I thank my marketing colleague at the Australian Graduate School of Management (AGSM), John Roberts, a recent MIT graduate, for alerting me to the Tournaments.

[2] It has recently been republished, with a new preface by Robert Solow (Marks 2018 and https://www.agsm.edu.au/bobm/papers/2018prefaces.pdf).

[3] From 1977 to 1982 I taught AGSM 713 Management in Society on the MBA Program. See the 1982 outline at https://www.agsm.edu.au/bobm/teaching/MinS82.pdf.

[4] Fourteen entries were submitted (by email) from three countries and five disciplines (Axelrod 1980a).

Axelrod announced the outcome of Tit for Tat in his first tournament and announced a second. Despite entrants' knowledge that Tit for Tat was very likely a competitor, no other submitted strategies outperformed it in the second tournament.[5] It was robust. But it is not always a winner: how well it does depends on the collection of competing algorithms in the round robin.

As is now widely known, Axelrod's tournaments revealed that one very simple strategy is robust in the IPD: Rapoport's Tit for Tat, which cooperates on the first round of the iterated game, and then mimics its opponent. Tit for Tat can be characterised as nice (start off cooperating), but easily provoked (defect after a single defect by its opponent), forgiving (a single cooperate by its opponent leads to its cooperating), and easily identified.

I found Axelrod's work fascinating, and, as an early user of the Internet myself,[6] I was intrigued by his use of it to research social interactions.

15.2 Generalising Axelrod's Iterated Prisoner's Dilemma Tournaments

In November 1984, MIT marketing professors Pete Fader and John Hauser decided to see whether Tit for Tat could be generalised in a three-person game with continuous actions. They ran one and then another competitive strategy tournaments (Fader and Hauser 1988). The Tournaments modelled competitors' decisions to price their outputs, which were imperfect substitutes: like a generalised Prisoner's Dilemma, without collusion among sellers, in a once-off interaction, pricing low (that is, defecting) is likely to be the outcome, at some cost to their profits, whereas pricing high (that is, cooperating) would result in higher profits for all, if no-one priced low.

I had been following Axelrod's tournaments and, given the chance, I submitted a FORTRAN algorithm for the First Tournament, but it was trigger-happy: if any of the players priced low (equivalent to defecting), then it would price low and keep pricing low. It was not quite Always Defect, but close to it.

The First Tournament was promising, but the organisers decided that it was flawed: their set-up had been too amenable to gaming, and so ran the Second Tournament with a different profit function (additive, not multiplicative[7]). Before the September 1986 deadline for this, John Roberts and I ran a local tournament, the AGSM Double Auction Computer Tournament, with a prize of $500, using the software from the Second MIT Tournament. The winner was Tony Haig (University

[5] There were 62 entrants from six countries and eight disciplines, recruited through announcements in journals for users of small computers (Axelrod 1980b).

[6] My web site has not changed in appearance at all since June 1996. The earliest record of my use of the Internet is ten years earlier, on 5 June 1986, at https://groups.google.com/forum/#!topic/net.text/rvi_FPmQBTE (Thirty-five years later I still use the Free Software Foundation's groff package for my text-processing needs: comes with all Macintoshes).

[7] Marks (1992b) presents these two profit functions.

of Western Australia). I then submitted an algorithm to the MIT Tournament; my algorithm won.

How did my algorithm win, people would ask. Some insights into the IPD, I would answer, but also luck: the ecology of the Tournament was comprised of all the other algorithms submitted; this meant that there was no easy way of coming up with a winning algorithm, especially since the other algorithms were unknown at the start. But, given that all algorithms were attempting to maximise their profits, a Nash outcome was likely.

The MIT Competitive Strategy Tournaments presented players with market conditions which generalised from the RPD. The environment resembled a more complicated version of the Iterated Prisoner's Dilemma, but there were more than two players, and there were more than two moves, since the algorithms chose a price in a range, while knowing the previous prices chosen by all players, but without collusion over the prices in this period.

As with the most productive research, there was an unanswered question: how had my algorithm won the 1987 MIT Tournament. It nagged at me.

15.3 Enter Machine Learning: The Genetic Algorithm

Mathematically, the problem of generating winning strategies in these interactions is equivalent to solving a multi-dimensional, non-linear optimisation with many local optima. In biological-evolution terms, it is equivalent to selecting for "fitness."

Indeed, in a footnote, Cohen and Axelrod (1984, p .40) suggest that:

> One possible solution may lie in employing an analogue of the adaptive process used in a pool of genes to become increasingly more fit in a complex environment. A promising effort to convert the main characteristics of this process to an heuristic algorithm is given by Holland (1975 [1992]). This algorithm has had some striking preliminary success in the heuristic exploration of arbitrary high dimensionality nonlinear functions.

In 1987 I read a colleague's copy of *Induction*, by Holland et al. (1986), which made a passing reference to some more recent work of Axelrod's which found strategies for playing the IPD Tournament which resembled Tit for Tat.[8]

Perhaps this work would shed light on my winning algorithm in the MIT Tournament. I emailed Axelrod and he replied with a program (written in Pascal-VS by Stephanie Forrest, his R.A.) using a version of John Holland's Genetic Algorithm, written in C. In an early example of a trans-Pacific code-sharing, I received the C

[8] I thank another AGSM colleague, applied psychologist Bob Wood, who brought the book back to Sydney from California and fortuitously lent it to me. John's and Bob's suggestions were very much helped by the cross-disciplinary atmosphere of the AGSM, sadly now lost after it was destroyed by incoming Vice Chancellor, Fred Hilmer.

code for the GA from the the U.S. Naval Research Laboratories in Washington DC and compiled it on our Unix machine.[9]

Mitchell and Forrest (1994) report what Axelrod did:

> Axelrod (1987) performed a series of experiments to see if a GA could evolve strategies to play [the IPD] game successfully. Strategies were encoded as look- up tables, with each entry (C or D) being the action to be taken given the outcomes of three previous turns.
>
> In Axelrod's first experiment, the evolving strategies were played against eight human-designed strategies, and the fitness of an evolving strategy was a weighted average of the scores against each of the eight fixed strategies. Most of the strategies that evolved were similar to TIT FOR TAT, having many of the properties that make TIT FOR TAT successful. Strikingly, the GA occasionally found strategies that scored substantially higher than TIT FOR TAT.
>
> To study the effects of a dynamic environment, Axelrod carried out another experiment in which the fitness was determined by allowing the strategies in the population to play with each other rather than with the fixed set of eight strategies. The environment changes from generation to generation because the strategies themselves are evolving. At each generation, each strategy played an IPD with the other members of the population, and its fitness was the average score over all these games. In this second set of experiments, Axelrod observed the following phenomenon. The GA initially evolves uncooperative strategies, because strategies that tend to cooperate early on do not find reciprocation among their fellow population members and thus tend to die out. But after about 10–20 generations, the trend starts to reverse: the GA discovers strategies that reciprocate cooperation and that punish defection (i.e., variants of TIT FOR TAT). These strategies do well with each other and are not completely defeated by other strategies, as were the initial cooperative strategies. The reciprocators score better than average, so they spread in the population, resulting in more and more cooperation and increasing fitness.

What I learnt was that, following up his own suggestion, Axelrod used Holland's Genetic Algorithm (GA) to "breed" strategies in the two-person IPD game (Axelrod 1987). He reported that the GA evolved strategy populations whose median member was just as successful as Tit for Tat, whom they closely resembled. I proceeded to teach myself Pascal-VS (I had learnt FORTRAN IV as an undergraduate and C later) and wrote a paper (Marks 1989a)[10] that attempted to replicate what Axelrod had done. In particular, I "bred" strategies playing in six distinct niches: (a) a niche of Always Defect, (b) a niche of Always Cooperate, (c) a niche of Tit for Tat, (d) a 5-rule niche described by Axelrod (1984, p. 199) that approximates his second IPD tournament, (e) an 8-rule niche from Axelrod (1987) that is a better approximation, and (f) the 5-rule Axelrod niche but with "noise" added, to simulate Nalebuff's (1987) IPD game with imperfect information. I experimented with strategies of different

---

[9] John Grefenstette's GENESIS 4.5. In doing so, I was probably the first economist in Australia to compile a C program obtained via the internet: I thank the first AGSM guru, Iain Johnstone, for choosing Unix as our OS.

[10] An earlier version of Marks (1989a) was presented at the 1988 Australasian meetings of the Econometric Society, Canberra, 30 August 1988, under the title, "Breeding hybrid strategies: the Prisoner's Dilemma computer tournaments revisited" (See *Econometrica* 57: 240, 1989, for the program).

complexities (depth of memory of past plays).[11] Marks (1989a) was the one of first papers presented by an economist.

using the genetic algorithm.[12]

The six niches in Marks (1989a) are static. It was suggested to me by David Schaffer (I think) that the this work would be more interesting if the strategies (agents) competed not against static niches but against themselves, as they evolved.[13] In Marks (1989b) I modelled this process.[14] What to call it? Unaware.

that the biologists had preceded me, I dubbed it "bootstrapping evolution;" but biologists (Ehrlich and Raven 1964) had called it "coevolution." My experiments showed that coevolution would result in convergence to cooperative behaviour in several strategic interations (the simple IPD, an extended IPD, and a simple version of the second MIT Competitive Strategy Tournament).

The genetic algorithm turned out to be a powerful computational method for searching for optima in spaces that were not amenable to calculus-based solutions, such as the oligopolistic pricing of the MIT Tournaments.[15]

With David Midgley and Lee Cooper, I wrote several more papers using the GA to examine strategies in various strategic interactions, usually market-related, and latterly with heterogeneous agents (with multi-population GAs) (Marks 1989b, 1992b, 2002a). We wrote several papers using the GA to explore optimal oligopolistic pricing (including Midgley et al. 1997, which, Shu-Heng Chen tells me, inspired Chen and Ni 2000).

An issue arose when I attempted to compare the weekly prices obtained from simulating the asymmetric agents derived using the GA with the historical data.[16] How to measure a distance between the dynamic historical data and different sets of simulated agents interacting? This can be thought of as a means of verifying the simulation models derived from the GA. Trying to answer this question has led to a series of papers and presentations, including Marks (2013) and culminating in Marks (2019). This has turned out to be a fruitful line of research.

So, from my win in the MIT Tournament, to my use of Holland's pioneering GA, to exploring historical oligopolistic reactions with David Midgley and other marketing colleagues, to the puzzle of verifying simulation output which is patterns (ordinal metrics) rather than simple (interval or ratio) metrics, I believe I have contributed to the theory and practice of simulation.

---

[11] Note that Tit for Tat is a one-round-memory strategy: no need for a deeper memory.

[12] Engineers had used the GA as a technique for optimising static functions. This was different.

[13] Another paper, Marks (1989b), was presented at the Allied Social Science Associations meetings under the auspices of the Econometric Society, on 30 December 1988, in New York. John Miller, a student at Michigan, wrote his Ph.D. on coevolution of automata playing the IPD; he published Miller (1986, 1989), and eventually Miller (1996). Other early economics papers to use GAs are Marimon et al. (1989), Miller and Andreoni (1990a, b), Holland and Miller (1991), Arifovic (1994).

[14] This paper was first presented at the North American Winter Meeting of the Econometric Society, New York, on 30 December 1988 (See *Econometrica* 57: 757, 1989, for the program).

[15] Oligopolies are markets with small numbers of sellers, each of whose actions in general affect the outcomes for other sellers, as well as for itself.

[16] See Marks (1992a) for discussion of the automata derived by the GA, and their simulations.

15.4 From Genetic Algorithms to Agent-Based Models

Another line of research has been agent-based (AB) models: When the individual agents modelled by the GA are competing against each other, the GA is modelling the process of coevolution. GAs were originally used as means of seeking optimal solutions to static problems; Marks (1989b) and others adapted them to seek solutions of coevolutionary strategic problems, such as the IPD and oligopolies with asymmetric players, where the fitness of an agent depends on the individual agents' actions, that is, the state of the whole population of agents.

When the interacting players face identical payoff sets and choose from identical action sets, a single population is satisfactory, since the GA processes that model learning among the individuals and between generations of the population are focused to the same end: faced with the same state of the interaction, any of the players would behave identically, and fitness is average (or discounted) profit.

A single population was acceptable when the players were not differentiated and when the flow of information from parents to offspring at the genotype level was not an issue, but when the players are modelling heterogeneous actors—in realistic coevolution, for instance—each player requires a separate population, not least to prevent the modelling of illegally collusive, extra-market transfers of information. This is discussed in Marks (2012).

For instance, Marks et al. (1999) develops an oligopolistic model with three (or four) interacting asymmetric sellers. The general model is still one of using a GA to search for automata (or mappings from past marketing actions to future actions for each seller, but now we use separate populations of agents, one population for each of the asymmetric sellers. This paper is one of the first to use the GA in an agent-based model (separate populations in the GA).

From separate populations of asymmetric agents, it is a simple step to develop agent-based models (ABMs). With David Midgley (see Midgley et al. 2007), we discuss AB models in marketing, specifically the complex interactions among three types of agents—consumers, retailers, and manufacturers—that lead to market and economic outcomes such as consumer satisfaction, and retailer and manufacturer profits. We argue that AB modelling is more appropriate than previous methods, but that it requires "assurance:" verification and validation of the model. The paper discusses this at length.

The work on AB models in economics led to an invitation from Peter McBurney and the editors of the journal to edit a special issue of *The Knowledge Engineering Review*, on agent-based computational economics. I asked Nick Vriend to be a co-editor (promising him little extra work), and we approached economists and financial economists to contribute. The special issue included papers from many of the pioneers in ACE.[17] One of the contributors was Shu- Heng Chen.

[17] June 2012; Editors: Robert Marks and Nick Vriend. Papers by: Anufriev and Hommes (2012), Arifivic and Ledyard (2012), Cheng Chan and Du (2012), Fagioli and Roventini (2012), Ladley (2012), Marks (2012), Page (2012), Richiardi (2012), Wilhite and Fong (2012).

From 1997 to 2010 I was the General Editor of the *Australian Journal of Management*, but, apart from three editorials (Marks 2002b, 2003, 2006b) and the invited papers Hailu and Schilizzi (2004) and Byde (2006), I did not focus on the topics of this chapter.

15.5 The Issue of the Best Risk Profile for Risky Decision Making

From my graduate studies, I had developed an interest in Decision Analysis, the use of mathematical tools to examine and prescribe decision making under uncertainty (Howard 1968). Given my publications using the GA, and my visits to the Santa Fe Institute in 1993 and earlier, in 1995 I was asked by John Casti to review a submission by George Szpiro to *Complexity* which used the GA to explore the best risk profile for a decision maker faced with uncertainty. To my MBA students, I had always taught that a slightly risk-averse profile was best: too risk-averse and the decision maker would pass up risky prospects ("nothing ventured, nothing gained"), but too risk preferring and the agent would sooner or later bankrupt itself.

After I gave him some guidance about how the GA found its solutions (not so frequently on the boundaries of the feasible space), Szpiro found that slightly risk averse was optimal, at least according to the simulation of the GA, and the paper was published (Szpiro 1997).[18]

But Szpiro had used an indirect method, despite my suggestion as referee that he explore using the GA to search more directly in the risk-profile space of utility functions.[19] Instead, Szpiro's agents are characterised by a variable (beta) which determines how many shares (in a simple market) they buy (or sell) in any period. For any model's parameters (see Szpiro's paper), there is a threshold dividend rate which determines whether agents should buy or sell. His GA model maximises each agent's wealth as a function of the dividend rate, and examines the evolved beta as uncertainty rises. He argues that betas that jump from zero to maximum at the threshold dividend rate reflect risk-neutral agents (and this occurs with no uncertainty), but that betas that respond to an increasing dividend rate more gradually (with uncertainty) from zero to maximum exhibit risk aversion. His agents cannot exhibit risk preferring, they are only risk neutral or risk averse. He argues that his model produces optimised agents whose risk aversion rises as uncertainty rises.

---

[18] Szpiro was later invited by Shu-Heng Chen to publish a chapter (Szpiro 2002) in the same volume that also included a chapter of mine (Marks 2002a). In later correspondence, after I had sent him a copy of Marks (2016), Szpiro told me that he was no longer involved in this line of research.

[19] Indeed, in his Footnote 6, Szpiro acknowledged my suggestion as referee to model agent automata as utility functions with explicit risk profiles, as an alternative to his approach (They could then, he allowed, be characterised as risk preferring, as well as risk neutral or risk averse).

I decided that one day I would follow my own advice and use the GA to try to confirm Szpiro's conclusion. I did try, but ultimately found a different result: risk neutrality.

Although I had previously (Marks 1989a and other papers) used a GA written in C, I decided to use the higher-level language NetLogo, with its support and graphic capabilities for the exploration of the best risk profile. Nigel Gilbert had implemented a genetic algorithm in NetLogo. The papers that followed presented the convergence results graphically.[20]

The title of the paper (Marks 2015a) I presented in London in 2014 was "Learning to be risk averse?" which reflected my persisting belief that Szpiro's results would hold with my more direct method of exploring the risk profile space for the best profile. Later that year I presented a revised paper (Marks 2015b) in Singapore, entitled "Searching for agents' best risk profiles," in which I concluded that for some utility functions (the wealth-dependent Constant Relative Risk Aversion function and the Dual-Risk-Profile function[21] from Prospect Theory) the best functions could be shown to be slightly risk averse, but for the Constant Absolute Risk Aversion function risk neutral was best.

In 2016, in a paper entitled "Risk neutral is best for risky decision making," (Marks 2016) I concluded that risk-neutral decision makers—whichever utility function was modelled—outperformed others when agents successively chose among three lotteries with randomly allocated probabilities and outcomes (two per lottery).[22]

Nonetheless, I realised that the simulation experiments in these papers were not very clear for readers, who might find the concept of the genetic algorithm difficult to be convinced by and the statistical arguments unfamiliar. That the GA is searching for an optimum risk coefficient at a flattish apex also clouds the findings. Furthermore, NetLogo has its uses, but it lacks a reputation for exact scientific work.

Although I had come to this line of enquiry by a quite independent route, as described above, I found that my path had crossed with Shu-Heng Chen's: his paper (Chen and Huang 2008) also looks at this question of the risk profile of decision makers under uncertainty; he had come to it from his work in developing agent-based artificial stock markets.

---

[20] I presented my evolving findings at the Complex Systems Research Summer School 2007 at Charles Sturt University in NSW and the 26th Australasian Economic Theory Workshop 2008 at Bond University in Queensland.

[21] For the DRP utility function, the risk profile is coded with two parameters, unlike the other two utility functions for which the search is in a single parameter space. See Sect. 3.3 in Marks (2020).

[22] The 2016 research used the NetLogo GA with the experimental setup: a population of 100 agents, each of which has a average winnings or a cumulative level of wealth, based on its risk profile and the successive outcomes of its 1000 choices among the lotteries; Each lottery is randomly constructed: the two payoffs ("prizes") are uniformly chosen in the interval [−\$100, +\$100], and the probability is chosen uniformly from [0,1]; Each agent faces 1000 lottery choices, and its cumulative winnings is that agent's "fitness" for the GA. The processes are stochastic. For each model we perform a number of Monte Carlo simulation runs to obtain sufficient data to analyse the results.

15.6 Simulation, Not Optimisation

Is risk neutrality best for making decisions under uncertainty? That is, for choosing amongst lotteries, for each of which the possible prizes and the probabilities of those outcomes were known? The analysis using the GA suggested risk neutral was best, but the results were not convincing. I decided I would search the risk-profile parameter space exhaustively, using simulation rather than searching the space using the GA. Perhaps not as elegant, but I hoped for a clear answer.

I decided not to continue using NetLogo (and certainly not to start coding in C); instead, I taught myself R.[23] Using R I wrote code to simulate exploration in the risk-profile space for eight methods of choosing among risky outcomes.[24] The paper that resulted (Marks 2020) demonstrates very clearly that risk-neutral decision makers outperform non-risk averse decision makers, whether CARA or CRRA or DRP utility functions (Marks 2020).

Two of the eight methods for choosing the best of the eight lotteries, max- min and max–max, are straightforward: max–min chooses the lottery with the highest minimum possible payoff, while max–max chooses the lottery with the highest maximum. Think of them as the pessimist's and optimist's methods, respectively. But the simulations revealed a surprising result which demands further research: the (pessimist's method) max–min, is almost twice as profitable as the (optimist's method) max–max. Just why is not yet clear. But even max–min is only a fraction as profitable as the risk-neutral Expected-Value method.

The question of which decision-making method gives the highest payoff in cases of uncertainty (where the possible pay-offs and their probabilities are known) is not, in general, amenable to closed-form solution. The answer is clearly that risk-neutral methods are best, as exemplified by the Expected Value method. I believe that exploration of other experiments in decision making under uncertainty (with complete information) will confirm the generality of this conclusion. Will relaxing our assumptions of complete information about possible outcomes and their probabilities result in different conclusions? This awaits further work. Marks (2020) was chosen for the Award for Best Paper at DECON 2019: the International Conference on Decision Economics, held in Ávila, Spain, on June 26−28, 2019.

---

[23] R is a programming language and free software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. It is widely used among statisticians and data miners for developing statistical software and for data analysis. As of January 2021, R ranks 9th in the TIOBE index, a measure of popularity of programming languages.

[24] See the R code at https://www.agsm.edu.au/bobm/papers/riskmethods.r. The R code sets up eight six-prize lotteries, with the prizes chosen randomly between +\$10 and −\$10. For each of eight methods of choosing one of the eight lotteries, there were 10,000 repetitions. See Marks (2020) for more details.

15.7 Connection with Shu-Heng Chen

I believe I first met Shu-Heng Chen at the 1st International Conference on Computational Economics, held at the University of Texas, Austin, May 21–24, 1995. At my invitation, in 2000 he came to Sydney and gave a presentation, "Genetic programming in the agent-based modeling of artificial stock markets," at the AGSM, UNSW, on February 15, 2000.

In 2001 he invited me to submit two papers for volumes that he was editing. The papers are Drake and Marks (2002) and Marks (2002a), published in *Genetic Algorithms and Genetic Programming in Computational Finance* and *Evolutionary Computation in Economics and Finance*, respectively. In return, some years later, as guest editor with Nick Vriend of the *The Knowledge Engineering Review*, a special issue on agent-based computational economics (ACE) (Marks and Vriend 2012),[25] I invited Professor Chen to contribute a paper, Chen et al. (2012).[26]

In July 2005 we were both in Bielefeld (at the the International Workshop on Agent-Based Models for Economic Policy Design, ACEPOL05) when Akira Nakatame invited some of us (including Shu-Heng and me, and other contributors to this volume) to become members of the Editorial Board of the new journal, the *Journal of Economic Interaction and Coordination*, the official journal of the new association, the Society of Economic Science with Heterogeneous Agents.[27] We have remained on the Editorial Board since then, and Shu-Heng is now co-editor.

At Professor Chen's invitation, in October 2005 I presented the Fourth Herbert Simon Seminar Series, on Agent-Based Computational Economics and Market Design, at the Artificial Intelligence Economics Research Center, Department of Economics, National Chengchi University, Taipei, and the National Kaohsiung University of Applied Sciences, Taiwan, on October 23–28.[28] These lectures were published as Marks (2006a), which, Shu-Heng Chen tells me, inspired his later research on agent-based modeling of lottery markets (Chen and chie 2008). At his invitation in 2013, I delivered three tutorials on "Validating Simulation Models, and Multi-agent Systems in the Social Sciences" at the Computational Finance and Economics Technical Committee (CFETC) of the Computational Intelligence Society (CIS) of the Institute of Electrical and Electronic Engineers, Inc. (IEEE) in Singapore in April 2013.[29]

---

[25] I thank Peter McBurney and the editors of the *KER* for this opportunity.

[26] With over 270 cites in Google Scholar, this is Professor Chen's second most cited paper.

[27] In November 2005 *JEIC*'s Editors in Chief: Akira Namatame, Tomas Lux, and Robert Axtell; Editorial Advisory Board: Mauro Gallegati, Masanao Aoki, and Alan Kirman; Editorial Board: Hideaki Aoyama, Yuji Aruka, Damien Challet, Shu-Heng Chen, Silvano Cinotti, Robin Cowan, Giorgio Fagiolo, David Green, Shouta Hattori, Dirk Helbing, Cars Hommes, Neil Johnson, Taisei Kaizoji, Sheri Markose, Matteo Marsili, Robert Marks, Denis Phan, Massimo Ricottilli, Erico Scalas, Frank Schweitzer, Didier Sornette, Hideki Takayasu, Leigh Tesfatsion, Zoltan Toroczkai, Bernard Walliser, David Wolpert, and Hirosi Yoshikawa.

[28] See my web page for these lectures, at https://www.agsm.edu.au/bobm/teaching/Taiwan.html.

[29] These tutorials are at https://www.agsm.edu.au/bobm/papers/IEEESingapore/tut01pr-3.pdf and https://www.agsm.edu.au/bobm/papers/IEEESingapore/tut02pr-3.pdf.

At Professor Chen's urging, I submitted a paper to DECON 2019: the International Conference on Decision Economics, held in Ávila, Spain, on June 26−28, 2019. That paper, Marks (2020), was judged the Best Paper at the Conference. Professor Chen was one of the conference organisers, and co-edited the conference proceedings.

Professor Chen's simulation study (Chen and Huang 2008) examines the survival dynamics of investors with different risk preferences in an agent-based, multi-asset, artificial stock market. They find that investors' survival is closely related to their risk preferences. Examining eight possible risk profiles, they find that only CRRA investors with relative risk aversion coefficients close to unity (that is, log-utility agents) survive in the long run (up to 500 simulations). This is not consistent with my findings in Marks (2020), whence I would expect risk-neutral agents to survive longer; these results remain to be reconciled.

My records reveal that Professor Chen and I have met at conferences in Austin, Texas (1995), Geneva, Switzerland (1996), Orlando, Florida (1999), Lake Arrowhead, California (2003), Kyoto (2004 and 2006), Bielefeld, Germany (2005 and 2010), Sydney (2009), Singapore (2013 and 2014), London (2014), and Ávila, Spain (2019).[30] Given our peripatetic movements (at least before COVID), there might well have been other occasions.

15.8 Conclusion

I have outlined how my research has evolved since 1987, the techniques I have used, and the results I have obtained. Moving from simple algorithms submitted to an in silico tournament of oligopolistic price competitors, to early adoption of the genetic algorithm in the Iterated Prisoner's Dilemma, and then to examining historical, asymmetric oligopolistic pricing decisions with multi-population agent- based models, and then to using the GA to search for the best risk profile for decision

[30] Respectively, the 1st International Conference on Computing in Economics and Finance (CEF 1995), Austin, May 21−24; the 2nd International Conference of Computing in Economics and Finance (CEF 1996), Geneva June 26−28; GECCO-99: the Genetic and Evolutionary Computation Conference, Orlando, July 13−17; the 2nd Lake Arrowhead Conference on Human Complex Systems, March 19−22; the 3rd International Workshop on Agent-based Approaches in Economic and Social Complex Systems (AESCS'04), Kyoto, May 27−29; the First World Congress on Social Simulation, Kyoto, August 21−25; the International Workshop on Agent-Based Models for Economic Policy Design (ACEPOL05), Bielefeld, June 30−July 2; and Advances in Agent-Based Computational Economics (ADACE 2010), Bielefeld, July 5−7; the 15th International Conference on Computing in Economics and Finance, UTS, July 15−17; Towards Large Multiscale Simulations of Complex Socio-Economic Systems of Heterogeneous Interacting Agents, the Society for Economic Sciences of Heterogeneous Interacting Agents (ESHIA), Nanyang, November 18−19; the 18th Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES'2014); the IEEE conference on Computational Intelligence for Finance Engineering & Economics (CIFEr'2014), London, March 27−28; and DECON 2019: the International Conference on Decision Economics, at the 17th International Conference on Practical Applications of Agents and Multi-agent Systems (PAAMS), Universidad de Salamanca, Ávila, Spain, June 26−28. As well as his presentation at UNSW Sydney in February 2000.

makers facing uncertainty, to exhaustive simulations of this issue, I have encountered my friend Shu-Heng Chen many times, and greatly benefitted from his advice, invitations, and suggestions.

If there is a moral for young researchers in all this, it is to follow one's nose or one's hunch. And listen to your students: several times students told me about research I was unaware of.[31] It also helps to talk with researchers in other disciples—marketing, computer science, political science, applied psychology—and I was fortunate to have such colleagues, including Shu-Heng Chen. The resulting papers might not appear in *Econometrica*, but it is now possible to publish them on-line, and there are also a growing number of cross-disciplinary outlets, as the Internet affects the core disciplines. Above all, use your imagination—technical skills are all very well (necessary) but hardly sufficient: a desire to answer previously unanswered questions (or perhaps to answer questions that have never previously been asked) requires a bold imagination.

References

Anufriev M, Hommes C (2012) Evolution of market heuristics. Knowl Eng Rev 27(2):255–271

Arifovic J (1994) Genetic algorithm learning and the cobweb model. J Econ Dyn Control 18:3–28

Arifovic J, Ledyard J (2012) Individual evolutionary learning with many agents. Knowl Eng Rev 27(2):239–254

Axelrod R (1980a) Effective choice in the Prisoner's Dilemma. J Conflict Resolut 24(1):3–25

Axelrod R (1980b) More effective choice in the Prisoner's dilemma. J Conflict Resolut 24(3):379–403

Axelrod R (1984) The evolution of cooperation. Basic Books, New York

Axelrod R (1987) The evolution of strategies in the iterated Prisoner's dilemma. In Davis LD (ed) Genetic algorithms and simulated annealing. Research notes in artificial intelligence. Morgan Kaufmann, Los Altos, CA, pp 32−41

Byde A (2006) Applying evolutionary search to a parametric family of auction mechanisms. Austr J Manag 31(1):1−16

Chen S-H (ed) (2002a) Evolutionary computation in economics and finance, studies in fuzziness and soft computing, vol 100, ed. by J Kacprzyk. Springer, New York

Chen S-H (ed) (2002b) Genetic algorithms and genetic programming in computational finance. Kluwer, Boston

Chen S-H, Chie B-C (2008) Lottery markets design, micro-structure, and macro-behavior: an ACE approach. J Econ Behav Organ 67(2):463–480

Chen S-H, Huang Y-C (2008) Risk preference, forecasting accuracy and survival dynamics: simulation based on a multi-asset agent-based artificial stock market. J Econ Behav Organ 67(3–4):702–717

Chen S-H, Ni C-C (2000) Simulating the ecology of oligopoly games with genetic algorithms. Knowl Inform Syst: Int J 2:310–339

Chen S-H, Chang C-L, Du Y-R (2012) Agent-based economic models and econometrics. Knowl Eng Rev 27(2):187–219

Cohen MD, Axelrod R (1984) Coping with complexity: the adaptive value of changing utility. Am Econ Rev 74:30–42

[31] Such as Hofstadter (1983), which is relevant here, and alerted many to the Axelrod tournaments.

Drake AE, Marks RE (2002) Genetic algorithms in economics and finance: forecasting stock market prices and foreign exchange—a review, in Chen (2002b), Chapter 2, pp 29−54. https://doi.org/10.1007/978-1-4615-0835-9_2.https://www.agsm.edu.au/bobm/papers/drake1.pdf
Ehrlich PR, Raven PH (1964) Butterflies and plants: a study in coevolution. Evolution 18:586−608
Fader PS, Hauser JR (1988) Implicit coalition in a generalised Prisoner's dilemma. J Conflict Resolut 32:553–582
Fagiolo G, Roventini A (2012) On the scientific status of economic policy: a tale of alternative paradigms. Knowl Eng Rev 27(2):163–185
Hailu A, Schilizzi S (2004) Are auctions more efficient than fixed price schemes when bidders learn? Aust J Manag 29:147–168
Hardin G (1968) The tragedy of the commons. Science 162(3859):1243–1248
Hofstadter D (1983) Computer tournaments of the Prisoner's Dilemma suggest how cooperation evolves. Sci Am, May 1983:6−26; reprinted in his Metamagical Themas: questing for the essence of mind and pattern. Basic Books, New York, pp 715−738
Holland JH, Miller JH (1991) Artificial adaptive agents in economic theory, Am Econ Revi 81:365–370
Holland JH (1992) Adaptation in natural and artificial systems. MIT Press, Cambridge, 2nd edn. (1975)
Holland JH, Holyoak KJ, Nisbett RE, Thagard PR (1986) Induction: processes of inference, learning, and discovery. MIT Press, Cambridge
Howard RA (1968) The foundations of decision analysis. IEEE Trans Syst Sci Cybern ssc-4:211−219
Ladley D (2012) Zero intelligence in economics and finance. Knowl Eng Rev 27(2):273–286
Marimon R, McGrattan E, Sargent TJ (1989) Money as a medium of exchange in an economy with artificially intelligent agents. J Econ Dyn Control 14:329–373
Marks RE (1989a) Niche strategies: the Prisoner's dilemma computer tournaments revisited, AGSM Working Paper 89–009 (April). An earlier version of this paper had been presented at the 1988 Australasian meetings of the Econometric Society, ANU, Canberra, 30 August 1988, under the title, "Breeding hybrid strategies: the Prisoner's Dilemma computer tournaments revisited" (See *Econometrica* 57:240, 1989, for the program). https://www.agsm.edu.au/bobm/papers/niche.pdf
Marks RE (1989b) Breeding hybrid strategies: optimal behavior for oligopolists. In: Schaffer JD (ed) Proceedings of the third international conference on genetic algorithms, George Mason University, June 4−7, 1989. Morgan Kaufmann Publisher, San Mateo, CA, pp 198−207. This paper had been presented at the Allied Social Science Associations meetings under the auspices of the Econometric Society, 30 December 1988, in New York (See *Econometrica* 57:757, 1989, for the program). https://www.agsm.edu.au/bobm/papers/1989P3ICGA.pdf
Marks RE (1992a) Repeated games and finite automata. In: Creedy J, Eichberger J, Borland J (eds) Recent developments in game theory. Edward Elgar, London, pp 43−64. https://www.agsm.edu.au/bobm/papers/creedychapter.pdf
Marks RE (1992b) Breeding hybrid strategies: optimal behaviour for oligopolists. J Evol Econ 2:17−38. https://doi.org/10.1007/BF01196459. An earlier version of this paper was presented at the Winter meetings of the American Economics Association in New York on 30 December 1988 under the auspices of the Econometric Society. https://www.agsm.edu.au/bobm/papers/JEE1992.pdf
Marks RE (1998) Competition and common propert. AGSM Working Paper 98−003 AGSM-UNSW Sydney, February 11. https://www.agsm.edu.au/bobm/papers/ccp.pdf
Marks RE (2002a) Playing games with genetic algorithms, Chapter 2, pp 31−44, in Chen (2002a). https://doi.org/10.1007/978-3-7908-1784-3_2. https://www.agsm.edu.au/bobm/papers/playing_games-shu.pdf
Marks RE (2002b) Editorial: simulating economics. Austr J Manag 27:i−vi, June. https://www.agsm.edu.au/bobm/editorials/0206edit.html
Marks RE (2003) Editorial: models rule. Austr J Manag 28:i−v, June. https://www.agsm.edu.au/bobm/editorials/0306edit.html

Marks RE (2006a) Market design using agent-based models, Chapter 27. In: Tesfatsion L, Judd KL (eds) Handbook of computational economics, Volume 2: Agent-based modeling. Elsevier Science, Amsterdam, pp 1339−1380. https://doi.org/10.1016/S1574-0021(05)02027-7.https://www.agsm.edu.au/bobm/teaching/SimSS/Marks.finalrev.pdf
Marks RE (2006b) Editorial: the threesome: TLA, BCP, and AFP. Austr J Manag 31:i−iv, June. https://www.agsm.edu.au/bobm/editorials/0606edit.html
Marks RE (2007) Validating simulation models: a general framework and four applied examples. Comput Econ 30(3):265−290, October. https://doi.org/10.1007/s10614-007-9101-7. https://www.agsm.edu.au/bobm/papers/s1.pdf
Marks RE (2012) Analysis and synthesis: multi-agent systems in the social sciences. Knowl Eng Rev 27(2):123−136. https://doi.org/10.1017/S0269888912000094. https://www.agsm.edu.au/bobm/papers/KER2012-pap.pdf
Marks RE (2013) Validation and model selection: three measures compared. Comp Econ 2:41−61, May. https://doi.org/10.7564/13-COEC10. https://www.agsm.edu.au/bobm/papers/parispaper.pdf
Marks RE (2015a) Learning to be risk averse? In: Serguieva A, Maringer D, Palade V, Almeida RJ (eds) Proceedings of the 2014 IEEE computational intelligence for finance engineering & economics (CIFEr), London, March 28−29. IEEE Computational Intelligence Society, USA, pp 1075−1079. https://doi.org/10.1109/CIFEr.2014.6924096. https://www.agsm.edu.au/bobm/papers/IEEEmarks2014.pdf
Marks RE (2015b) Searching for agents' best risk profiles. In: Handa H, Ishibuchi M, Ong Y-S, Tan K-C (eds) The proceedings of the 18th Asia Pacific symposium on intelligent and evolutionary systems (IES'2014), Chapter 24, vol 1. Springer, pp. 297−309. https://doi.org/10.1007/978-3-319-13359-1_24. https://www.agsm.edu.au/bobm/papers/marksIES2014.pdf
Marks RE (2016) Risk neutral is best for risky decision making, mimeo, 16 March 2016. https://www.agsm.edu.au/bobm/papers/rnisbest.pdf
Marks RE (2018) Non-renewable resources and disequilibrium macrodynamics. Routledge, London (Reprinted version of my 1978 Stanford dissertation, with a new Preface by Robert Solow.) (Routledge Library Editions: Environmental and Natural Resource Economics, vol 9)
Marks RE (2019) Validation metrics: a case for pattern-based methods. In: Beisbart C, Saam NJ (eds) Computer simulation validation.: Springer International Publisher, Cham, Chapter 13,pp 319−338. https://doi.org/10.1007/978-3-319-70766-2_13. https://www.agsm.edu.au/bobm/papers/ch12_marks_rev4.pdf
Marks RE (2020) Calibrating methods for decision making under uncertainty. In: Bucciarelli E, Chen S-H, Corchado JM (eds) Decision economics: complexity of decisions and decisions for complexity. Springer, pp 1−9. https://doi.org/10.1007/978-3-030-38227-8_1. https://www.agsm.edu.au/bobm/papers/marksspain.pdf
Marks RE, Vriend NJ (2012) The special issue: agent-based computational economics—overview. Knowl Eng Rev 27(2):115−122. https://doi.org/10.1017/S0269888912000082. https://www.agsm.edu.au/bobm/papers/00marksvriend.pdf
Marks RE, Midgley DF, Cooper LG, Shiraz GM (1999) Coevolution with the genetic algorithm: application to repeated differentiated oligopolies. In: Banzhaf W, Daida J, Eiben AE, Garzon MH, Honavar V, Jakiela M, Smith RE (eds) GECCO-99: Proceedings of the genetic and evolutionary computation conference, July 13–17, 1999, Orlando, Florida. Morgan Kaufmann, San Francisco, pp 1609−1615. https://www.agsm.edu.au/bobm/papers/Gecco1999-RW-766.pdf
Midgley DF, Marks RE, Cooper LG (1997) Breeding competitive strategies. Manag Sci 43(3):257−275, March. https://doi.org/10.1287/mnsc.43.3.257. https://www.agsm.edu.au/bobm/papers/MS-MMC97.pdf
Midgley DF, Marks RE, Kunchamwar D (2007) The building and assurance of agent-based models: an example and challenge to the field. J Bus Res 60(8):884−893, August. https://doi.org/10.1016/j.jbusres.2007.02.004. https://www.agsm.edu.au/bobm/papers/jbr-mmk.pdf
Miller JH (1986) A genetic model of adaptive economic behavior, University of Michigan Working paper

Miller JH (1989) The coevolution of automata in the repeated prisoner's dilemma. Working paper 89-003 Santa Fe Institute
Miller JH (1996) The coevolution of automata in the repeated Prisoner's dilemma. J Econ Behav Organ 29:87–112
Miller JH, Andreoni J (1990a) A coevolutionary model of free riding behavior: replicator dynamics as an explanation of the experimental results. Working paper at Santa Fe Institute and University of Wisconsin
Miller JH, Andreoni J (1990b) Auctions with adaptive artificially intelligent agents. Working paper 90-01-004, Santa Fe Institute
Mitchell M, Forrest S (1994) Genetic algorithms and artificial life. Artif Life 1(3):267–289
Nalebuff B (1987) Economic puzzles: noisy prisoners, Manhattan locations, and more. J Econ Perspect 1:185–191
Page SE (2012) Aggregation inagent-based models of economies. Knowl Eng Rev 27(2):151–162
Richiardi MG (2012) Agent-based computational economics: a short introduction. Knowl Eng Rev 27(2):137–149
Schelling TC (1978) Micromotives and macrobehavior. Norton, New York
Szpiro GG (1997) The emergence of risk aversion. Complexity 2:31–39
Szpiro GG (2002) Tinkering with genetic algorithms: forecasting and data mining in finance and economics, Chapter 15, pp 273−286, in Chen (2002)
Tesfatsion L (2021) Agent-based computational economics: overview and brief history, Chapter in this volume
Wilhite A, Fong EA (2012) Agent-based models and hypothesis testing: an example of innovation and organizational networks. Knowl Eng Rev 27(2):221–238

Chapter 16
Extending Herbert Simon's "Science of Design": the Role of Collaboration and Users in Development of Technically Advanced Systems

Ben Vermeulen and Andreas Pyka

Abstract While Herbert Simon's formalistic perspective on the process of design is widely influential, it has shortcomings in prescribing the development of technically advanced systems. Notably, it does not explicate how the development requires decomposition into subsystems and subsequent distribution of tasks over domain experts. In addition, it has an immutable "outer environment" spanning requirements and thus does not address how development iterates between articulation of user requirements and technical specification co-evolving towards a concrete design. Given the distributed nature of development activities, expert agents working on specific problems and subcomponents need ways to overthrow the current decomposition and project organization to escape unforeseen development lock-ins, technical infeasibility, and market unviability. This chapter extends Simon's "science of design" with his notions of problem-solving heuristics and system modularization and synthesizes a conceptual process model of system decomposition, organization of collaborative development, and exploration of user requirements. Moreover, it stipulates the role of communication and decision heuristics for agents working concurrently and consecutively in a collaborative, distributed development process.

Keywords Technology development · Design process · Collaboration · Market uncertainty

16.1 Introduction

In studying conceptual formalizations of the process of development of systems, one is bound to run into Herbert Simon's work. Herbert Simon introduced a wide variety of ground-breaking concepts to describe the (ideal–typical) process

B. Vermeulen (✉)
IQIB, Bad Neuenahr-Ahrweiler, Germany
e-mail: ben.vermeulen@iqib.de

A. Pyka
Institute of Economics, University of Hohenheim, Stuttgart, Germany

R. Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-15294-8_16

of searching for a design (Simon 1988), recursively solving problems (Simon and Newell 1971), discovering laws of nature (Simon 1977, 1992), etc. Moreover, he provided concepts and insights in the properties of the structure of systems, notably near-decomposability (Simon 1962). As a testimony to the applicability, his concise conceptual models of learning, cognitive problem-solving, and design processes have laid the foundation for artificial intelligence, expert systems, and other types of decision support software. Indeed, Simon wrote, elatedly, that this "science of design" is not only possible, but is actually emerging in engineering schools, notably in computer science and systems engineering, and in business schools through management science (Simon 1988). As such, one may expect that Simon's work may also be applicable in describing the (ideal–typical) process of development of technically advanced systems. Optimistically, a Simonian formalistic approach could ultimately form the basis for design support software or even unattended artificial intelligence coming up with designs for such systems.

Despite the optimistic outlook, we argue that the "science of design of the artificial" falls short for real-world design processes, notably when it comes to technically advanced systems. Just as in the theory of human problem-solving, Simon essentially focused on the tasks/ activities of *a single focal* information processing system applying heuristics to adapt an 'inner environment' to the immutable 'outer environment'. A wide variety of scholars have readily elaborated on limitations of the stylizations of Simon (Visser 2010; Hatchuel 2002; Ohlsson 2012; Dorst and Dijkhuis 1995; Carroll 2006). For one, Hatchuel (2002) argues that the process of design is characterized by an ill-structured nature and a formalization thereof should as such allow unexpected expansion of concepts and constraints. Notably, the design process also includes distinguishing what is the 'inner' and 'outer' environment, learning what is to be learned, and even designing learning devices to do so. A particularly blatant omission as source of uncertainty is the relevant presence of other agents and the role of social dynamics (see e.g. Carroll 2006). After all, particularly for technically advanced systems, over the course of materialization there are problems to be solved and design decisions to be made that are not within the technical expertise of the focal agent. As such, the system may thus get decomposed into subsystems, and certain developments may be allocated to different domain experts. These domain experts subsequently collaboratively develop the system at hand. Moreover, the organization of the development project may well evolve with the technical challenges to be solved, particularly if the targeted functionality or design changes. Given the potentially complex interaction across subsystems in the different domains, there is substantial uncertainty on the feasibility during the progressive materialization. Moreover, particularly when (sub)systems are tested in real-world settings, the potential users of the system and other stakeholders may have different requirements than foreseen and may thus trigger a reset of the design and introduce new technical problems to be solved.

This chapter pays heed to both the factors of uncertainty and particularly the 'social side' of the development process. Note that, when it comes to design processes, there are more holistic, less formal approaches, such as design thinking (see the review in Micheli et al. 2019), and rather more structured methods such as new product development (Ulrich and Eppinger 2016; Cooper 2004). These methods do actually incorporate 'the social side' and allow adapting to unexpected change of concepts and constraints. However, unlike Simon's science of design, these methodological frameworks are conceptually quite involved. Instead, here, we revisit Herbert Simon's framework to introduce collaboration of developers as well as customers (or rather, henceforth: users) in the invention, development, and design of an advanced system. In particular, we extend Simon's theoretical concept of iterative tree traversal for technical system development with several own concepts. These concepts are to highlight the co-evolutionary interaction of the system development process with search for the organization of the project with various experts and market exploration. We hereby build on insights from two previous studies: one on the process of invention of the heavier-than-air aircraft (Vermeulen and Guffarth 2017) and an extensive ethnographic study on the design of robots (Sorenson et al. 2019).

The structure of the chapter is as follows. A stylistic model of the process of technology development is given in Sect. 16.2. Generally, in the development of technically advanced systems (such e.g., aircraft and robots), agents with different expertise are involved. Section 16.3 describes the collaborative development process. Moreover, quite unlike other studies on collaboration in design, problem-solving, or development, a more long-term evolutionary perspective is taken. Development of many systems starts off with functional decompositions to -in part- be able to rely on 'off-the-shelf' components made by others. As such, the concept of 'collaboration' can be stretched to describe a process of accumulation, decomposition, selection of designs, artifacts, shelved knowledge, etc. fragmented across space and time *possibly (but not necessarily) without working together explicitly*. In addition, in developing a technical system, the developers may well have to discover who to design for. So, another obvious but quite different agent central in any technology development process is the user (which Simon occasionally referred to as "the client"). The role of users is described in Sect. 16.4. Although users are less explicit in early invention processes that typically focus on mere technical feasibility, the requirements of users are best not ignored in the design and development of products for a commercial market. Not only may user requirements be uncertain, they may evolve and depend on design decisions. Moreover, there may be different users with requirements that cannot be reconciled technically in one and the same design. A synthesis of Simon's work on design science and our additions on the role of collaboration and users is provided in (Sect. 16.5). From the framework thus obtained, it is clear that developers will have to prudently bootstrap out of the market and technological uncertainty by having (potential) users articulate requirements based on pilots with experimental technology. This has to be done, in conjunction, while making design and system decomposition decisions to be able to divide and distribute development tasks over agents collaborated with, which subsequently impede system changes and constrain markets that can be served.

16.2 A Process Perspective of Technology Development

Here, we present a conceptual perspective on the process of 'technology development', with a particular interest in creating a technically advanced system that is new to the world. In our view, technology development includes activities ranging from coming up with concepts, imagining and constructing the physical form, testing in real-world settings, experimenting with technical alternative elements to further advance particular properties, etc. The process of developing a complex, technically advanced system (e.g., aircraft, robot) includes finding a configuration of components that provides particular functionality to the envisaged user of the system. Technology development encompasses (i) invention (leading to a rudimentary but functioning instance of something new to the world), (ii) innovation (enhancing the performance of the system in real-world applications), and (iii) design (imagining, conceptualizing, or shaping the instance/ physical artifact following a narrative of its use and functioning). The perception of what entails 'designing' differs between academic fields and communities of practitioners, but a more detailed treatise is out of scope of this chapter Note that an invention is a concrete, physical artifact and an innovation is a further enhancement/ improvement thereof, while a design may be a physical instance, but also just a plan or narrative of the functioning of the invention or innovation. So, as such, designing is a more generic activity. The technical feasibility of a design can only be established by testing the artifact in the real world. In development and design of a technically advanced system, developers are constantly engaged in (i) heuristic search for solutions to technical problems; (ii) recursing through task trees and iterating between system design, subsystem development, and testing; and (iii) bootstrapping out of market and technical uncertainty. These three activities are now described in more detail.

16.2.1 Search Process, Including the System Design

Extending the formalistic perspective of search processes in problem-solving and scientific discoveries so characteristic of his work, Simon argues that design is searching through a space of possible artifacts, thereby using heuristics and learning instruments. To make use of Simon's work as the basis for our description of the ideal–typical process of developing a technically advanced system, we revisit two central conceptions. Firstly, following Simon, we perceive the system as a hierarchy of subsystems. Decomposing a complex system in the form of hierarchies allows making search a near-decomposable process (Simon 1962): "*Subparts belonging to different parts only interact in an aggregative fashion – the detail of their interaction can be ignored*". Secondly, the process of coming to a design for that system is also regarded as a problem-solving activity. Hereby, the problem-solver is plagued by bounded rationality and thus needs to resort to heuristics. Unlike the critique of some design methodologists (Dorst and Dijkhuis 1995), Simon did not argue design was a rational optimization process, but rather that it is boundedly rational search following a certain logical regularity. Instead, due to a variety of cognitive limitations (such as working memory restrictions, computational power), human developers resort to rules-of-thumb, effort reduction mechanism, etc. generally plotting non-optimal, heuristic search trajectories through design problem space. Heuristics are applied to move from solution to solution for a technical problem, from design to design, and decomposition to decomposition for a system. Depending on the evaluation criteria, there are some selection mechanisms that condition the structure of the system. In fact, natural selection drives the evolutionary emergence of near-decomposability (see e.g., Simon 1962). As such, from here on, it is assumed that the system is a near-decomposable hierarchy of subsystems. There is a wide variety of advantages, not only from the perspective of system performance, but also in the design and technical problem-solving at the level of the subsystems. Such a structure facilitates the organization of the project and distribution of tasks across agents with relevant expertise (see Sect. 16.4).

16.2.2 Recursing Through Task Trees

There are various ways in which new technology comes into existence, but whether it is by gradual materialization of abstract concepts or from combining existing artifacts and evaluating the use, development involves solving technical problems. Adopting Simon's formalization of problem-solving activities, technology development is perceived as traversing through task trees and recursively solving technical bottlenecks that in turn also require formulating (new) low-level task trees to subsequently traverse. However, while doing so, inventors may encounter (yet) unresolvable or insurmountable issues, which thus feed back to higher levels and may cancel and revise parts of the current task tree. Technology development thus involves recursively solving technical problems to ultimately come up with an instance for experimentation, testing, and reflection on choices. The results of these tests may in turn also lead to redefining goals, representations, and defining new (sub)trees. Moreover, the task tree itself is based on the system decomposition and development tasks for each subsystem. Given that the process of development for each subsystem may thus take place across time and space, developers have to cope with requirements or specifications that have been fixed before. Alternatively, the developers may fix them under uncertainty about the technical feasibility or market viability. Upon encountering issues, the developer has to decide whether to develop something afresh, overthrow the existing requirements, whether to abandon existing methods, look for alternative tools, etc.

Following existing complexity-theoretical conceptualizations given in Vermeulen and Guffarth (2017), technology development up and until invention is an iterative, recursive process consisting of

(i) Defining functionality ('what should it do'), formulating a technological decomposition into interlocking subsystems jointly providing particular functionality ('how to do that');
(ii) Designing individual subsystems, experimenting in restricted/ laboratory settings, thus gaining an understanding of operational principles, which in turn possibly lead to a reformulation of system design, functionality or configuration;
(iii) Assembling the full system for tests in different configurations under real-world circumstances, which may lead to a redevelopment or even a complete redefinition of the system being invented.

This model will be extended in the remainder of this chapter.

16.2.3 Bootstrapping Out of Uncertainty

Since technology development is characterized by technological uncertainty ('unknown unknowns'), the developer is likely to encounter unforeseen technical issues that lead to updates of the task tree. In addition, the exact system requirements often only become clear with further materialization and subsequent expounding in practice with potential users. Changes in the requirements to fulfill may also lead to additional tasks to fulfil. However, who is to be considered the targeted user often only becomes clear with successful materialization. As such, there is a 'dilemma of sequentiality': one has to decide to either rely on ill-defined demand to guide developments that may be hard, costly, or impossible to realize or instead develop experimental technology for a market that may not exist in the end. Developers now have to sequentially 'bootstrap' out of the market and technological uncertainty by gradually and alternatingly fixing more of the market requirements or fixing more of the design and material form. For instance, whenever an experimental product is tested with potential users, it may not yet fulfil its function fully (e.g., tests with a construction co-bot revealed that it was yet too slow and unwieldy for the workers) or may not be suitable for each user and thus exclude some unintentionally (e.g., tests with an educational robot revealed that it could not be properly operated by schoolchildren because the controls were made for the bigger hands of adults). Picking a particular application may subsequently restrict the material form or change the concept, which limits the users to target (e.g., a harvest robot requires a leveled guide rail, making it suitable for modern greenhouses but leaving it unsuitable for variable outdoor situations). These findings would subsequently trigger new development activities or rather change the market to target. As such, arguably, the task tree is to contain a mix of activities in obtaining market requirements and technology development. Given this bootstrapping, there is co-evolution of fixed technical specifications and articulation of user requirements. In addition, developers may conduct develop-test-plan cycles with users but may decide to do in a 'staggered fashion', i.e. only gradually extend the set of stakeholders involved over time (see Sect. 16.4).

16.3 Collaborative Development

Simon focused on the tasks and activities of *a single* information processing agent. Hereby, this agent applies heuristics in designing an artifact mostly from the ground up, seeks to discover laws and regularities, and solves problems arising. In this section that conceptual framework is extended with notions of collaboration. The development of a technically advanced system often requires collaboration of a multitude of specialized experts. In some cases, the development of such a system takes place in a single team and is organized and coordinated methodically in one common project. In many cases for a mass-produced product, a supply chain of firms (e.g., automobile, cellphone) produces components that are assembled by a system-integrator. In this case, there is incremental development of otherwise fairly standardized components in Research & Development departments of vertically specialized firms. In other cases, the development of a yet ill-defined system takes place haphazardly in different projects across time and space, often without any formal links, with technological knowledge shelved in books or passed down embodied in artifacts, etc. However, in general, development projects require search and problem-solving in both the design and the organization space in conjunction. Here, two elements of the process of collaborative development are discussed. Firstly, there is accumulation of technology fragmented across time and space, which underlines the significance of taking stock of what other agents are doing. Secondly, there is a challenging relationship between (i) system design and its decomposition, (ii) the decoupling between and matching of development tasks and skills of agents in the organization of development, and (iii) options to change design or reverse development decisions given the progressive organization of industries. These two elements are discussed below.

16.3.1 The "Blackboard" with Fragmented Accumulation of Technology

In general sense, technology accumulates and evolves over time (Basalla 1988), whereby discovery of one may lead to the discovery of another (see e.g., Arthur 2014). As such, quite naturally, developers often do not know the origins of general technological knowledge, tools, etc. However, even in concrete projects, when asked about the origin of design decisions, resources and artifacts used, and other agents involved, developers are often only able to reveal a small part (Sorenson et al. 2019). In many cases, design considerations are embodied in artifacts passed down, knowledge is shelved as book, possibly codified by actors not or no longer involved, and in many cases information on decisions is missing the context. In fact, the course of technology developments in the narrow context of a concrete project may be affected quite strongly by what is known, available, etc. to those developing.

An analytical lens on technology accumulation is Simon's "blackboard", which is a long-term memory used to communicate information across different subtasks regardless of where in the task tree the agent is working (Simon 1977, p.297). The use of such a blackboard is a long-standing, commonly used technique in distributed problem-solving (see e.g., Decker 1987). Here, it is perceived as (possibly partial) memory to share information across time and space on imaginaries, intermediate (re)design considerations, reports on failed and successful tests, experimental subsystems, artifacts, even search heuristics, etc.

To illustrate the fragmented accumulation of technology and significance of information stored on the blackboard, let's look at the (well-documented) process of invention of the heavier-than-air aircraft ultimately culminating into the Wright brothers' aircraft. For an elaborate description, the reader is referred to Vermeulen and Guffarth (2017). In this case, a whole range of inventors developed imaginaries, concepts, toys, artifacts, and even test facilities, fragmented over the entire nineteenth century in several countries (UK, France, Germany, USA, notably). In many cases, these inventors built upon prior work done in different eras, in different places; often without direct communication; this pertains to coming up with a functional design (e.g. ornithopter), system decomposition (e.g., the combination of wings for lift and a propeller for thrust), or a subsystem that is (re)used (e.g., the steam engine). It was already for George Cayley in 1809 to formulate a configuration of a fixed wing, a fuselage to carry the pilot, and an adjustable horizontal and vertical tail for stabilization and control. William Henson read his work and designed and received a patent on his Aerial Steam Carriage in 1843, of which technical drawing and images of an aircraft with propellers flying (e.g., over the Egyptian Pyramids) appeared in newspapers internationally. Only few inventors largely ignored practical design considerations of others before. There are counterexamples, but these actually underline the value of taking stock of what is on the blackboard. Clement Ader, for instance, spent considerable amounts of money on developing his bat-like Eole and Avion in the 1880s and 1890s. However, the designs were already considered ridiculous in that era. The Wright brothers argued that Ader, by his "*slavish imitation of nature*", had produced wings with not enough ribs to give sufficient support or uniform lift, ultimately making "*the whole machine [] ridiculous*" (Hallion 2003, p.136–137). However, on the other hand, taking stock of and building upon existing technology may actually also inhibit progress. Particularly the maturity of the technology may have developers (erroneously) focus on other subsystems as bottleneck. For instance, with steam engines driving trains and busses in many countries, the first experimental heavier-than-air aircraft were also equipped with steam engines (e.g. Henson & Stringfellow's Ariel in 1843 and triplane in 1868 in the U.K., Thomas Moy's Aerial Steamer in 1875, Maxim's vehicle in 1890, Ader's Eole in 1890 in France, but also Langley's Aerodrome models in the U.S.A. in the early 1890s). However, these engines simply had a highly unfavorable power-to-weight ratio.

This notion of fragmented accumulation does not concern just components alone, but also very generic skills (e.g., the practical experience in making artifacts such as train parts or bicycles), availability of certain tools (e.g., the wind tunnel), or the overall functional design (having to have lift, overcome drag, provide thrust). Some of the developments were deliberate following up on efforts of others. For instance, the Wright brothers read on the use of wind tunnels in Chanute's 1884 book. They proceeded building one in 1901, following examples of Wenham in 1871 and Phillips in the 1880s. Moreover, while the Wright brothers are credited for the significance of their focus on control of aircraft rather than autostability, they developed the inferior mechanism of 'wing warping' for roll control, instead of using Boulton's ailerons invented in 1868 or Goupil's elevons invented in 1883. Arguably, efforts of Samuel Langley (his publications and study visits to Europe) and particularly Octave Chanute (his many articles and 1894 book "Progress in Flying Machines" containing findings of many inventors, the 1893 conference he organized, his extensive correspondence with the Wright brothers, etc.) were of great significance as 'shelved technological knowledge'.

The notion of fragmented accumulation and the role of the blackboard therein underlines that, regardless of whether these agents are actively communicating, the technology development is distributed and quasi-collaborative. The significance of this 'blackboard' of course speaks in favor of patents to disclose technological knowledge. In addition, there are various concepts in which agents (mostly firms) freely share their design and development results for other agents to extend, despite potentially competing interests, see e.g. 'collective invention' (Allen 1983; Nuvolari 2004; Powell and Giannella 2010).

16.3.2 Organization of Collaborative Development

Apart from the *implicit* "collaboration" by means of using technological knowledge accumulated on a collective blackboard, there is often *explicit* collaboration. During the last couple of decades, a variety of organizational concepts for collaborative development have emerged, such as distributed innovation, open innovation, collaborative innovation, private collective innovation, etc. These conceptual frameworks focus mostly on the organizational modes, legal matters, business models, etc. Since the early 80s, however, there is also mounting attention for low-level formalistic frameworks on collaboration., e.g. in problem-solving (e.g., Cammarata et al. 1983; Smith and Davis 1988), design (e.g., Fischer et al. 1992; Cross and Cross 1995; Decker 1987), innovation (e.g., Lakhani and Panetta 2007; Baldwin 2012), and decision-making (e.g., Bonabeau 2009). The collaborative development of a technically advanced system may require formulating subtasks and distributing these over domain experts. The solutions to the various subproblems/ tasks can then ultimately be synthesized into one (see, e.g., Smith and Davis 1988; Cammarata et al. 1983).

In general, collaborative development has vertically specialized agents work in parallel and in relative autonomy on their individual development subtasks. Ideally, these subtasks are allocated based on the match with their skills, expertise, knowledge, and resources. The fact that collaborative development often has agents work in relative autonomy on their subsystem of course poses new challenges. The design choice, system decomposition, and subsequent task allocation in the early phases 'feed forward' into the development options for the agents in later phases. These specialized agents then work within the search space allocated to them, but may have an incomplete and possibly even incorrect representation of the outer environment or possibly limited information on actions of other agents. The technological solutions for subsystems used in the current design are then largely based on decisions taken by others, often already earlier. As we have seen in Sect. 16.2, the actual development and design process is iterative and recursive (at least, at the engineering level). Particularly whenever design choices have to be made, a burst of communication between agents would have to be accommodated. The design of complex, technically advanced systems often involves long-term, indirect communication on the design process itself, knowledge of artifacts created by others, and rationales for previous design decisions— all of which calls for supporting systems with long-term memory (Fischer et al. 1992).

Whenever industries start to emerge, modularization of designs (i.e., reducing technical relationships by standardization of interfaces between modules) facilitates distribution of development tasks between vertically specialized agents (cf. Baldwin and Clark 2000; Frenken 2006; Sanchez and Mahoney 1996). As a consequence of the limitations in the design space for agents both in individual projects or in industrial development, there are organizational impediments of system change. If there is vertical specialization, developers responsible for different subsystems are employed by different firms and institutes. The ability to (give direction to) change the system design are then limited by the boundaries of firms and institutes and their relationships. In fact, it is a severe challenge for an individual agent to reverse design and development decisions across subsystems. So, arguably, upon encountering a bottleneck during the traversal of the task tree, it should be possible to collectively change the system design and the organization of the development activities. In absence thereof, developments may be largely locked into the system design from the start.

The division, distribution, and allocation of the subtasks as well as communication procedures are actually part of the problem (Cammarata et al. 1983). Indeed, the group of agents needs to decide and agree on the organization and communication policies to deal with matters such as sharing of resources and resolving interdependencies between subtasks across technical interfaces (see e.g., Cammarata et al. 1983). On top of the purely organizational challenges to changing the system design, agents may also have conflicting strategic interests. For example, agents may seek to use technologies that are not compatible. Some form of arbitration may be needed to resolve such conflicts and move forward with the technology development.

16.4 Coping with Users in the 'Outer Environment'

Recent history is littered with experimental products that, once piloted in a real-world context, proved to be technologically inadequate, excluded particular users unintentionally, left users concerned about safety or privacy, etc. Many of these products never diffused on a large scale and only few of them are now still remembered. However, technologies such as LaserDisc, Windows 8, MySpace, Google Glass, the Concorde, and the Blackberry phone enjoyed a modest or brief market uptake but were soon outcompeted by alternative designs. Many aspects of the design are decided upon rather early on in the development process. However, consequences of these decisions may become clear only during tests, pilots, or even only upon implementations. While 'the client' ultimately is the provider of 'needs and wishes' to be targeted, the science of design does acknowledge that the 'outer environment' may be unclear and in fact evolve, e.g. subject to changing professional norms and government interventions (Simon 1998, p.150). Moreover, the actual 'outer environment' is actively shaped; explicit demand is often non-existent and partly created. So, there is uncertainty about who is 'the' user and, moreover, these users have ill-defined requirements. As such, it makes sense to discuss when, why, and how developers (should) make particular design decisions, and when, why, and how users are (to be) involved. Early involvement of users is at the forefront of design methodologies such as new product development and design (Ulrich and Eppinger 2016; Cooper 2004; Cooper and Kleinschmidt 1995), co-creation and co-design (Sanders and Stappers 2008; Prahalad and Ramaswamy 2004), and design thinking (Micheli et al. 2019). Seemingly, it is commendable to follow a -what we refer to as the- micro-level "develop-test-plan cycle" strategy of observing users engage with early designs and then plan new development activities (Sorenson et al. 2019). However, such early involvement may not be a guarantee for "success". There are several reasons as to why artifacts may not be technically feasible or market viable (yet) and why early involvement of users may actually be challenging. Here, it is elaborated how developers are posed with the challenges that (i) the requirements of users may not be known and may even be irreconcilable, (ii) the requirements depend on the materialization and vice versa (and yet design decisions have to be made at some point), (iii) other stakeholders such as intermediaries and representatives are also present and introduce their own (biased?) requirements, and lastly (iv) development decisions affect options to engage with users and hence adapt the design later.

16.4.1 Irreconcilable Requirements of Users

The development of an artifact commences with (implicit) articulation of its functions for targeted users. In case the size of the various market segments and the preferences of the users are known (and constant), marketing research models would theoretically allow optimization of product designs (cf. Green and Krieger 1989). However, there is substantial market uncertainty, as not only are there different user segments of which the size may be unknown (rather than a single 'representative' user), also their exact requirements need not be known yet. While well-possibly deemed out of scope of Simon's perspective, a method for practical design of artifacts should arguably deal with the uncertainties about requirements on functions and usage. Given that near-exhaustive exploration of all user requirements would be costly, developers resort to satisficing and halt once a 'sufficiently clear' image of those requirements is obtained. However, it is arguably not only about not knowing requirements or the exact distribution of numbers of users with particular requirements. Reconciling the requirements of the various segments in a single design may be commercially unfavorable or technically challenging. It may well be that developers have to make trade-offs (Consumer X likes A and B, but the design can technically not offer both at the same time), make in-/exclusion decisions (Consumer X likes A, Consumer Y dislikes A), and resolve design conflicts (Consumer X likes A, Consumer Y likes B, but the design cannot offer both at the same time).

Given the user diversity and possible 'irreconcilability' of hard functional requirements, the developer is to make choices about design aspects in relationship to the users that are thereby in/excluded. Sorenson et al. (2019) provides the description of a case in which a co-bot is developed to lift, transport, and keep in place heavy objects during construction. The robot was saluted particularly by a few older workers, which, after having it done for many years, dreaded the 'backbreaking task' with these heavy objects. However, the robot was at that time still quite slow and unwieldy, and younger workers preferred doing the task manually as they could thus earn more under the piece-rate pay. Given that teams were composed of both such older and younger workers, the robot ended up not being used at all. Of course, from the designers' perspective, such trade-offs, conflicts, and exclusions would ideally be discovered early on, so as to prevent consuming time or resources on exclusive design priorities. To this end, designers should expound their designs in an explicit user context. One way to resolve irreconcilable requirements may be to modularizing the subsystems that inadvertently exclude certain types of users. The developer may then develop a subsystem variant for each type of user.

While there are algorithms that optimize product design and market segment choices (cf. Green and Krieger 1989), these require knowing the market size and preferences. While such algorithms are expected to perform poorly under uncertainty, they may perform relatively well under unbiased noisy observations of the users' preferences and requirements. To obtain these observations, one should interview potential users or even 'crowd source' the actual selection of system features and functions to broad panels of users. This may be a viable strategy for generic technology with obvious functions and for which there are many potential users (e.g. household appliances). It is less so for technically advanced systems with specific users and involved functions (e.g. professionals, institutions). Vermeulen et al. (2020) reports on experiments with human subjects with a serious game on product design and market exploration under resource scarcity. The experiments revealed that 'unfocused' exploration of market segmentation, each segment's requirements, as well as available product technologies significantly increased chances of designing a technically feasible and market viable product. Moreover, this study also revealed that this early exploration increased the likelihood that subjects actually find the largest market segment for which it can develop feasible product technology.

16.4.2 Bootstrapping in specifications and requirements

In Simon's perspective, the 'inner environment' is adapted to the 'outer environment', i.e. more specifically, the system design is adapted to requirements of real-world users. However, in addition to not knowing these requirements or possibly not being able to reconcile them in the design, the requirements may also become only apparent when the system is largely developed and used during trials or even after implementation. Indeed, the typical 'linear' or 'waterfall' engineering approach is to assume particular user requirements and characteristics of the environment, develop an experimental system, and then engage in adaptation and finetuning after running pilots in (staged) real-world settings. Apart from having an incomplete or even wrong impression of the requirements of users initially, it may also be that the artifact is later used in a context that was not considered early on in the design. One may think of different types of users (e.g. gender, age, handedness), use in a different environment (e.g. outdoor vs indoors), or use in a different cultural context (e.g. a Korean voice-operated robot used in Finland). Designs may thus have elementary flaws such as the use of language or symbols (e.g. on buttons), appearance (e.g. toy-like versus looking 'complicated'), etc.

However, developers need to either first fix the direction to develop technology and then pick the market to enter and users to target or first pick the market to then focus on developing the technology meeting the requirements of users in that market. In the first case, the developers may be convinced that "*people don't know what they want until you show it to them*" (quoting Steve Jobs). While this may occasionally be true, there is also a certain risk. Once a trial version of the system is developed and expenses made, tests with users may reveal that initial design decisions were flawed ('started developing the wrong technology') or that there is actually no demand for the product. Developers may work with imaginaries on what the *supposed* users require, but the (real) users may have different 'needs and wishes'. Conversely, selecting particular potential users and rigidly pursuing their 'needs and wants' is also risky. After all, development activities may stray away from existing technological expertise ('wrong' targets), may lead to (unnecessarily) costly developments, may cause squandering resources for market segments ultimately not targeted, or the developers may end up with feasible technology but for an ultimately commercially unattractive niche. The solution to this 'dilemma of sequentiality' is to bootstrap out by alternating between obtaining user feedback with increasingly more specific designs and trying to develop new technology based on increasingly more articulated user requirements. One thus gets a temporal interleaving of market and technology research with a gradual convergence toward a more concrete system and more specific market segment to target. Moreover, particularly when investments in development activities are more specific for targeted market segments or require a substantial part of the total budget available, one may want larger samples of potential users. This would give more certainty on the market size and requirements.

16.4.3 Representatives, Intermediaries and Not Just Users

While determining the market segment and fixing the requirements is already challenging for many systems, there often are also other stakeholders besides the users, such as intermediaries and representatives. One may think of the physiotherapist and family members also using technology that essentially supports health care patients, or parents and teachers also using educational tools developed for children (Sorenson et al. 2019). A way out of the aforementioned 'dilemma of sequentiality' is also a strategy of 'staggered expansion' of the group of stakeholders to be involved in develop-test-plan cycles. The developers may first act purely on the basis of assumptions about the user, then involve intermediaries (possibly in several iterations), and in the later stages start to fine-tune with the final users (possibly in several iterations). A word of warning, though, as obtaining user requirements and information on the environment of application from those other than the user allows for blind spots, biases, and ignorance of these intermediaries to creep into the design.

16.4.4 Testing of Subsystems

Uncovering practical issues or additional low-level user requirements is also complicated due to decomposition of a system into subsystems and distribution of development tasks over domain experts. Developments in each subsystem may be very impactful for the performance for the entire system, yet testing with users may not be possible until the entire system is assembled. Consequently, it may well be possible that users can only be involved at stages in which impactful decomposition and functionality decisions have already been taken. This becomes all the more obvious when thinking of advanced socio-technical systems to be operated in a human-centered service setting (e.g. a hospital, construction site) or become part of a larger system (e.g. farm, warehouse) in which the community of designers and developers at large has yet little experience. In this case, it may well happen that, in the test or implementation stages, the system will be used by non-professional users that give unexpected input or use the output unexpectedly, etc. Such 'interactions' may well be so idiosyncratic that they are only uncovered in test trials, pilots, or even actual use after implementation.

16.5 Synthesis

In the previous sections, it was shown how Simon's theory of an ideal–typical search process in technology space is to be complemented with search in the project organization space and the market space in conjunction. At the project-level, there is

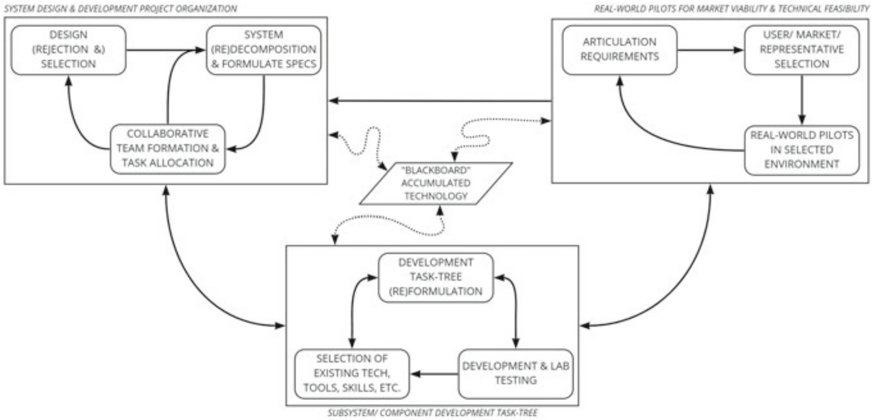


Fig. 16.1 Process model of technology development

search for a design, its functionality definitions, and the decomposition into subsystems. In addition, in general, a team of agents has to be formed, and these agents are to be allocated to development tasks for subsystems. At the level of (sub)systems, developers either select and use existing technologies, tools, etc. or decide to develop such subsystems afresh and formulate, traverse, and update task trees, all interleaved with laboratory tests. Subsequently, tests with subsystems in isolation or the entire assembled system are conducted in real-world pilots to allow (i) further articulation of user requirements, updating the targeted user market, and (ii) determining technical issues to be resolved at the subsystem level. These tests may lead to resetting technical specifications or rejection/ reformulation of the design with targeted functionality. A synthetic process model of technology development learning loops in the style of Simon is given in Fig. 16.1.

Design Selection concerns adopting a conceptual design, possibly based on existing technology or analogies with nature, but definitely building upon imaginaries of the functionality ('what should it do'). System Decomposition & Formulation of Technical Specification concerns breaking down the system into subsystems to provide certain functionality. Developers have to resolve the 'dilemma of sequentiality' by alternating between obtaining demand information as guide for development under uncertainty about technological feasibility or developing experimental technology based on (premature) user requirement for which there may as such not be a market in the end. Developers prevent irreversible investments by interleaving technology developments with exploring and acquiring user requirements.

The Team Formation and Task Allocations concerns how, based on the subsystems and components to be developed, a team of agents is formed and tasks are allocated to the team members, based on the expertise, skills, etc. The search for, selection, and matching of agents is based on developments required. The composition of the team and activity of members may change over time, depending on what is needed and planned. However, this decomposition and team formation has to be done rather early on, when user requirements are possibly not yet fully known and cannot be yet acquired because the system does not yet exist. The choice for subsystems may be based on existing, accumulated technology, standard methods, etc., which may form a source of lock-in, inhibiting development. Moreover, the (subteams of) agents developing subsystems may reside at different firms or institutes and are thus confined by the boundaries thereof.

The Development and Lab Testing concerns how individual agents or subteams are engaged in invention, design, and development of subsystems. Agents recursively go through development task trees that are being updated based on test results and findings stored on the Blackboard. This blackboard plays a pivotal role in design and development decisions, as it contains generic collections of books (shelved codified knowledge), artifacts (embodied knowledge), imaginaries, as well as highly specific technical decisions made locally. From this blackboard, developers obtain information collected across time and space and thus prevent duplication, make informed decisions, and focus on bottlenecks in performance. It is also used to inform team members and allow sensible decisions in formulating and traversing the development task tree in new iterations. Findings may lead to a reformulation of system design and functionality, and thereby even of the composition of the group of agents and tasks allocated to these agents. The division of tasks and distribution over different agents may complicate and thereby make reversing design and development decisions difficult and costly.

Developments are to lead to Real-World Pilots in which subsystems and eventually the full assembly are tested in real-world circumstances with real target users. This may lead to discovery of technical issues or articulation of new or alternative requirements. This may trigger redevelopment of components or even a complete redefinition of the system being invented. The Articulation of User Requirements may be consolidating, refining, or rather reveal issues with the technical specifications assumed for development. Whenever the user requirements are irreconcilable, developers will have to make decisions on which users to in- or exclude, which technical features to provide and which not, or decide on a trade-off. In order to bootstrap out of the situation with market and technological uncertainty, developers may seek to expand the group of stakeholders in a staggered fashion. So, first start with internal assumptions, then involve representatives and intermediaries to obtain provisional (and possibly biased) requirements, and only later get more detailed requirements from potential users with concrete experimental designs.

16.6 Conclusion

The goal of this chapter was to study the adequacy of and extend Simon's "science of design" framework for development of complex, technically advanced systems. The development and design of a system is seen as boundedly rational, heuristic search through design space, thereby traversing development task trees and using the blackboard to communicate across development stages. One serious shortcoming in

Simon's framework is that there is no 'social side'. In his framework, a single agent engages in development, design, problem-solving, etc. alone. However, in many present-day projects in practice, systems are developed in collaboration: other agents perform some development tasks and solve subproblems in their expert domain. Even the supposedly lone inventors build upon accumulated technological knowledge produced by other developers scattered across space and time. Collaborative development however introduces serious challenges of its own because the distribution and allocation of tasks across domain experts is based on the premature decomposition of the system design. To overcome system lock-in, these expert agents thus need ways to overthrow the current decomposition. Ideally, then, other agents are concurrently involved instead of in a previous phase. Another shortcoming in Simon's work is that the 'outer environment' is partly spanned by user requirements and should be actively searched as well. Given that there is fundamental market *and* technological uncertainty, developers have to overcome a 'dilemma of sequentiality'. Notably, they have to 'bootstrap' out by alternating between, on the one hand, finding potential users and obtaining user feedback with increasingly more specific designs and, on the other hand, further developing technology based on increasingly more articulated user requirements.

Simon's framework for the search for feasible system designs has thus been extended with search for the organization of the collaborative development project and obtaining user requirements through market exploration *in conjunction.* This is then not done by one and the same agent anymore, but by different agents fragmented across space and time and each with limited observations and partly serving their own interests. As such, new communication and decision heuristics are to be introduced to attune these search processes across agents.

Another 'social side' of the outer environment that Simon ignored and we have not been able to discuss in this chapter is the role of competition. Simon did not seek to explain the drivers behind the accumulation of the vast portfolio of artifacts and the motives of agents to engage in technology development. For Joseph Schumpeter, in contrast, it is exactly the existence of autonomous agents competing for scarce market returns that drives technology development and explains many of the design choices made. After all, entrepreneurs on the supply side develop different designs and the customers on the demand side deselect designs that have inferior properties. Standardization and modularization may be a sensible strategic management approach, as it would allow outsourcing component production to other agents. This in turn triggers component innovation of competing upstream producers, reaping scale advantages of producing components for multiple customers, etc. Moreover, agents change the market conditions ('outer environment'), e.g. by launching a cheaper, equivalent product or a superior product in higher demand. Such a Simonian-Schumpeterian perspective on technology development is developed in, e.g. Chie and Chen (2013), Chen and Chie (2004), Chie and Chen (2014), Vermeulen and Pyka (2014a), Vermeulen and Pyka (2014b) and Vermeulen et al. (2017).

Acknowledgements This chapter builds upon work done in several projects and Ben Vermeulen gratefully acknowledges funding for the DFG DACH grant PY 70/8-1 / 214145179; DAAD project

57219050; H2020 project 731726; and NWO grant 458-03-112. Anecdotal evidence and examples on aircraft come from Vermeulen and Guffarth (2017) and on robots from Sorenson et al. (2019).

References

Allen RC (1983) Collective invention. J Econ Behav Org 4(4):1–24

Arthur WB (2014) The nature of technology. What it is and how it evolves. Free Press

Baldwin CY (2012) Organization design for distributed innovation. SSRN J. https://doi.org/10.2139/ssrn.2055814

Baldwin C, Clark K (2000) Design rules: the power of modularity. MIT Press

Basalla G (1988) The evolution of technology. Cambridge University Press

Bonabeau E (2009) Decisions 2.0: The power of collective intelligence. MITS Loan Manag Rev 50(2)

Cammarata S, McArthur D, Steeb R (1983) Strategies of cooperation in distributed problem solving (Rand Note, N-2031-ARPA)

Carroll JM (2006) Dimensions of participation in simon's design. Des Issues 22(2):3–18

Chen S-H, Chie B-T (2004) Functional modularity in the fundamentals of economic theory: toward an agent-based economic modeling of the evolution of technology. Int J Mod Phy B 18(17):2376–2386

Chie B-T, Chen S-H (2013) Non-price competition in a modular economy: an agent-based computational model. Econ Polit: J Anal Inst Econ 33(3):273–300

Chie B-T, Chen S-H (2014) Competition in a new industrial economy: toward an agent-based economic model of modularity. Admin Sci 4(3):192–218

Cooper RG (2004) New Products—What separates the winners from the losers and what drives success. In: Kahn KB (ed) The PDMA handbook of new product development, 2nd edn. Wiley, New York, pp 3–28

Cooper RG, Kleinschmidt EJ (1995) Benchmarking the firm's critical success factors in new product development. J Prod Innov Manag 12(5):374–391. https://doi.org/10.1111/1540-5885.1250374

Cross N, Cross A (1995) Observations of teamwork and social processes in design. Desig Stud 16(2):143–170

Decker KS (1987) Distributed problem-solving techniques: a survey. IEEE Trans Syst, Man, Cyber 17(5):729–740. https://doi.org/10.1109/TSMC.1987.6499280

Dorst K, Dijkhuis J (1995) Comparing paradigms for describing design activity. Design Stud 16(2):261–274

Fischer G, Grudin J, Lemke A, McCall R, Ostwald J, Reeves B, Shipman F (1992) Supporting indirect collaborative design with integrated knowledge-based design environments. Hum-Comput Interact 7(3):281–314. https://doi.org/10.1207/s15327051hci0703_2

Frenken K (2006) Innovation, evolution and complexity theory. Edward Elgar, Cheltenham

Green PE, Krieger AM (1989) Recent contributions to optimal product positioning and buyer segmentation. Eur J Oper Res 41(2):127–141. https://doi.org/10.1016/0377-2217(89)90375-5

Hallion R (2003) Taking flight. Inventing the aerial age from antiquity through the First World War. Oxford University Press, Oxford

Hatchuel A (2002): Design theory: the unachieved program of Herbert Simon. In J Manage Govern 5 (3–4).

Lakhani KR, Panetta JA (2007) The Principles of distributed innovation. Innov Technol, Govern, Global 2(3):97–112. https://doi.org/10.1162/itgg.2007.2.3.97

Micheli P, Wilner SJS, Bhatti SH, Mura M, Beverland MB (2019) Doing design thinking: conceptual review, synthesis, and research agenda. J Prod Innov Manag 36(2):124–148. https://doi.org/10.1111/jpim.12466

Nuvolari A (2004) Collective invention during the British industrial revolution: the case of the Cornish pumping engine. Camb J Econ 28:347–363
Ohlsson S (2012) The problems with problem solving: reflections on the rise, current status, and possible future of a cognitive research paradigm. J Probl Solving 5(1). https://doi.org/10.7771/1932-6246.1144
Powell WW, Giannella E (2010) Collective invention and inventor networks. In: Handbook of the economics of innovation, vol 1. Elsevier (Handbook of the economics of innovation), pp 575–605
Prahalad CK, Ramaswamy V (2004) Co-creation experiences: the next practice in value creation. J Interact Market 18(3):5–14. https://doi.org/10.1002/dir.20015
Sanchez R, Mahoney JT (1996) Modularity, flexibility, and knowledge management in product and organization design. Strat Mgmt J 17(S2):63–76. https://doi.org/10.1002/smj.4250171107
Sanders EB-N, Stappers PJ (2008) Co-creation and the new landscapes of design. CoDesign 4(1):5–18. https://doi.org/10.1080/15710880701875068
Simon HA (1962) The architecture of complexity. Proc Ame Philos Soc 106(6):467–482
Simon HA (1988) The science of design: creating the artificial. Des Issues 4(1/2):67–82
Simon HA (1998) The sciences of the artificial, 3rd edn. MIT Press, Cambridge
Simon HA, Newell A (1971) Human problem solving. The state of the theory in 1970. Am Psycho 26(2):145–159
Simon HA (1977) Models of discovery. And other topics in the methods of science. Dordrecht: Springer Netherlands (Boston Studies in the Philosophy of Science, 54). https://books.google.nl/books?id=iCvpCAAAQBAJ
Simon HA(1992) Scientific discovery as problem solving. Int Stud Philos Sci 6(1)
Smith RG, Davis R(1988) Frameworks for cooperation in distributed problem solving. In: Readings in distributed artificial intelligence. Elsevier, pp 61–70
Sorenson J, Zawieska K, Vermeulen B, Madsen S, Trentemøller S, Pyka A et al (2019) Perspectives on robots. A reality check on imagined futures. Aarhus University, Copenhagen
Ulrich K, Eppinger SD (2016) Product design and development, 6th edn. McGraw-Hill, New York
Vermeulen B, Guffarth D(2017) A process model of invention and the role of government, Institutions, and geography. anecdotal evidence from the aerospace industry in the years 1800–1950. In: Vermeulen B, Paier M (eds) Innovation networks for regional development. Concepts, case studies, and agent-based models, 1st ed. Springer International Publishing, Cham (Economic Complexity and Evolution), pp 97–129
Vermeulen B, Pyka A (2014a) Technological progress and effects of (Supra) regional innovation and production collaboration. an agent-based model simulation study. In: IEEE conference on computational 3/27/2014a - 3/28/2014a, pp 357–364
Vermeulen B, Pyka A (2014b) The effects of supraregional innovation and production collaboration on technology development in a multiregional world: a spatial agent-based model study. In Was J, Sirakoulis G, Bandini S (eds) Cellular automata. Lecture notes in computer science, vol 8751, pp 698–707
Vermeulen B, Chie B.-T, Chen S.-H, Pyka A (2017) Evolutionary programming of product design policies. An agent-based model study. In: IEEE Staff (ed) 2017 21st Asia Pacific symposium on intelligent and evolutionary systems (IES). Hanoi, 11/15/2017 - 11/17/2017. IEEE, Piscataway, pp 1–6
Vermeulen B, Chie B-T, Pyka A, Chen S-H (2020) Coping with bounded rationality, uncertainty, and scarcity in product development decisions: experimental research. In: Bucciarelli E, Chen S-H, Corchado JM (eds) Decision economics: complexity of decisions and decisions for complexity. Advances in intelligent systems and computing, vol 1009. Springer International Publishing, Cham, pp 24–35
Visser W(2010) Simon: design as a problem-solving activity. Art Design & Psych (2):11–16