UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Heidelberg University

Alfred Weber Institute of Economics

Dissertation:

Understanding and Modeling Economic Behavior: Experimental Insights and Computational Perspectives

A Thesis Submitted to Obtain the Degree of:
Doctor of Philosophy (Ph.D.)

Presented by:

Vinícius Ferraz

Supervisor: Prof. Dr. Christiane Schwieren

March, 2024

Dissertation

Submitted to the

Alfred Weber Institute of Economics

Heidelberg University, Germany

for the degree of

Doctor of Philosophy (Ph.D.)

Presented by

Vinícius Ferraz

Supervisor

Prof. Dr. Christiane Schwieren

Subject

Understanding and Modeling Economic Behavior: Experimental Insights and Computational Perspectives

---

Oral examination: March, 2024

Abstract

In the multidisciplinary field of economic behavior, traditional theories often struggle to capture the inherently complex nature of human decision-making processes. Building upon well-established theoretical foundations, this research proposes a comprehensive exploration of human economic behavior. It leverages the strengths of behavioral economics, experimental methodologies, and advanced computational techniques, integrating these into comprehensive analytical models. Through five interconnected papers, the core objective is to investigate the psychological complexities of individual choices. Rigorous experimental designs and robust methodologies reveal detailed insights into actions and decisions. The introduced studies cover adaptive and evolutionary learning, equilibria in asymmetric games, as well as fairness and loss aversion in strategic interactions. They also investigate the correlation between dark personality traits and dishonesty, explore the phenomenon of algorithm aversion, and examine the dynamics of motivated sampling of information. These topics collectively provide a broad perspective on human decision-making in economic contexts, with the findings offering deep insights into diverse real-world inspired scenarios. To achieve this, the research utilizes advanced computational techniques such as Genetic Algorithms, Agent-Based Models, Reinforcement Learning, Machine Learning, and Causal Inference. From understanding the psychological mechanisms underlying decision-making to examining well-established behavioral traits like loss aversion and dark personality traits, this dissertation paints a comprehensive picture. It adeptly bridges the gap between theoretical constructs and real-world implications, presenting a fresh perspective on the dynamic nature of economic behavior in contemporary society.

Acknowledgements

First and foremost, I'd like to express my deep gratitude to Prof. Dr. Christiane Schwieren for the great opportunity and privilege to embark on this incredible adventure. I am grateful for the chance to become part of the scientific community and pursue my goals and ideas, always with enthusiastic support and excellent mentorship. Thank you for believing in me from the beginning.

Prof. Dr. Thomas Pitz has been indispensable in this journey. I thank you immensely for the long-lasting collaboration, the teachings, and the consistently excellent supervision and advice. Equally important, I thank Prof. Dr. Jörn Sickmann for his advice and support. Both of you enabled a significant portion of this research project. I thank Prof. Dr. Rosemarie Nagel for the always kind and expert feedback, for sharing the depths of her knowledge, and also for being part of this thesis committee.

I am thankful to my co-authors and research colleagues, Leon, Tamás, Cosima, Deniz, and Wolf. Collaborating with each of you has been insightful, pleasant, and rewarding. Your support and perspectives have undoubtedly enriched this work.

A special mention to Dr. Serhan Ili and my work colleagues: sharing this academic journey with you has added depth to the experience. I appreciate the shared insights, support, and camaraderie.

A thank you to friends and family who closely follow this journey: Carolina, Samuel, Gregory, Marcel, and Glendi.

Finally, this thesis is dedicated to my loving family: Marcia, Marcio, Edison, Victor, and Enzo. Thank you all for the unconditional support, love, and guidance.

Contents

List of Figures

List of Tables

1. Introduction

Economic behavior, rooted in the complex tapestry of human decision-making, presents a multifaceted challenge to traditional economic theories. As individuals navigate the complexities of economic scenarios, their choices often diverge from classical rational, self-maximizing predictions. This divergence has paved the way for the emergence of behavioral economics, which seeks to understand economic decisions' psychological, cognitive, and emotional underpinnings. The field has leaned heavily on experimental methods to empirically validate these behavioral insights, employing controlled settings to test and refine theories. In parallel, the digital age has enabled unprecedented data availability and an ever-evolving collection of computational methods, offering a glimpse into real-world economic behaviors on an unimaginable scale. This confluence of behavioral insights, experimental rigor, and computational prowess presents challenges and opportunities. It allows researchers to harness advanced analytical tools and bridge the gap between controlled experiments and the complex realities of human actions in the real world.

This research project is designed to comprehensively explore economic behavior through various scenarios, analytical lenses, and facets. Building upon prior research, the aim is to uncover fresh insights that shed light on the complexities of human decision-making. This endeavor is articulated through five papers, each sticking to the same foundational principles. These papers navigate diverse contexts, each presenting unique dynamics and focusing on decisions influenced by elements other than rationality per se.

From a contextual perspective, Behavioral economics, influenced by elements of psychology, challenges the traditional economic theory that assumes individuals are purely rational and self-maximizing. As Thaler (2016) suggests, this field provides both normative and descriptive models that yield better predictive power for real-world scenarios. Over the past decades, evidence has mounted, indicating that human decision-making is governed by heuristics and biases and is influenced by factors such as the choice environment, incentive structures, and information presentation (Tversky & Kahneman, 1974). These individual-level behavioral anomalies can have ripple effects, influencing entire market structures, as Shiller points out Shiller (2003). The traditional notion of the *homo economicus* is consistently challenged, emphasizing the need to bridge everyday economic life with models of individual choice behavior (Henrich et al., 2001).

Parallel to behavioral economics, experimental methods form the next pillar of this research. As defined by V. L. Smith (1991), experimental economics is a discipline that studies human behavior in controlled settings, allowing the assessment of theories and models. It offers a platform to test hypotheses about decision-making, isolating specific factors to understand their effects. Samuelson (2005)'s work further highlights the trajectory of experimental economics, evolving from a niche area to a cornerstone of economic research. The literature draws attention to the alignment of experimental findings with traditional economic theories and the complexities that arise when theoretical predictions diverge from practical outcomes.

The synergy between behavioral and experimental economics is profound. While experimental economics offers controlled environments to test and validate theories, behavioral economics goes deeper into decision-making's cognitive and emotional facets. This combination ensures that experimental settings test and validate behavioral theories, leading to a comprehensive understanding of economic behavior. The interconnection between these fields offers a broad perspective, bridging theoretical insights with real-world economic decision-making, a connection discussed by Santos (2011).

In recent advancements within the field, the application of machine learning and other computational resources in behavioral and experimental economics has gained traction. Mullainathan and Spiess (2017) emphasize the shift from procedural to empirical approaches akin to econometrics. They note that machine learning's strength lies in its ability to predict and uncover generalizable patterns from complex data structures. This predictive power, distinct from traditional economic parameter estimation, offers a fresh perspective and tools for understanding economic behaviors. Building on this perspective, C. F. Camerer (2018) underscores the utility of machine learning in behavioral economics, emphasizing its adeptness at harnessing a broad spectrum of variables for predictive purposes. Machine learning's ability to sift through extensive variable sets and pinpoint predictive ones aligns seamlessly with behavioral economics objectives, offering a richer understanding and a promising avenue for subsequent research.

Athey (2018) also underscores the transformative potential of machine learning in economics, highlighting its capability to enhance policy decisions, estimate causal effects, and address generalization concerns in economic contexts. Green, White, et al. (2023) share that machine learning techniques offer robust methods for causal inference and characterizing treatment effect heterogeneity in experiments. Furthermore, the importance of understanding the differences between machine learning and traditional econometrics is emphasized in Athey and Imbens (2019), suggesting that machine learning techniques require adaptation to address specific economic challenges. The integration of these tools is advocated as essential for the empirical economist's toolkit, ensuring a comprehensive approach to economic research.

Complementarily, Plonsky et al. (2019) highlight the pivotal role of behavioral decision theories in predicting human choices in economic decision tasks. Their findings underscore the synergy between machine learning and behavioral insights, emphasizing that integrating qualitative behavioral tendencies and quantitative descriptive models enhances predictive accuracy in human economic behavior. Similarly, I. Lundberg, Brand, and Jeon (2022) discuss the significant impact of machine learning in advancing social science research. They highlight its ability to enhance traditional methods, facilitate more profound discoveries, and integrate human reasoning with computational prowess. The authors champion the collaborative approach of human expertise and algorithmic automation, emphasizing that while machine learning offers powerful tools, it complements rather than replaces the human element in research.

Recent developments in the literature suggest an evolution in the landscape of economic research. As such, it becomes perceptibly imperative to approach the study of economic behavior from multifaceted perspectives. Following this reasoning, this research is anchored on three foundational pillars: behavioral economics, experimental methodologies, and computational techniques. While distinct in its approach, each pillar collectively contributes to a comprehensive understanding of economic decision-making. Behavioral economics offers insights into the psychological complexity of individual choices, experimental methodologies provide the empirical evidence to validate or challenge existing theories, and computational techniques equip us with the tools to analyze and interpret complex datasets. Together, these pillars form the basis of my exploration into economic behavior, as detailed in the subsequent sections.

Regarding the understanding and modeling of economic behavior, a sentence that sets the title of this manuscript, the focus is on understanding the psychological mechanisms underlying decision-making. With the indispensable support of my co-authors, I aim to elucidate how individuals make choices, considering potential biases and influencing factors. My examination of specific traits such as loss aversion, inequity aversion, dark personality traits, algorithm aversion, and motivated decision-making highlights this investigation. To empirically validate the formulated hypotheses, I employ rigorous experimental methodologies. By conducting laboratory, online, and simulated experiments with varied treatments and designs, I gather data that provides insights into individual behavior in different strategic scenarios. To further analyze and interpret this data, I leverage computational methods. My research incorporates advanced computational techniques, including Genetic Algorithms, Random Forests, Gradient Boosting Machines, Causal Machine Learning, Agent-Based Modeling, and Reinforcement Learning. These methodologies enable us to process complex data relationships and derive findings that might be challenging to obtain through traditional analytical methods. In essence, by leveraging computer simulations, intricate experimental designs, and a blend of conventional statistical analyses with advanced machine learning and causal methods, I aim to offer profound insights and methodological advancements to the broader field of economics.

1.1 Motivation and Objectives

The combination of behavioral, experimental, and computational economics provides a holistic approach to understanding the intricacies of decision-making and strategic behavior. With the ever-advancing nature of science, the role of technology in reshaping our understanding of economics becomes increasingly evident. Therefore, the motivation behind the research conducted in this project's scope takes inspiration from these four pillars.

Technological Advancements in Computational Models: The rapid advances in computational technologies have revolutionized our ability to process vast amounts of data. This allows for efficiently handling large datasets and facilitates uncovering intricate interactions and patterns previously obscured.

Methodological Advancements: There is a growing aspiration to push the boundaries of traditional methodologies in economics. By integrating computer science methods into experimental and behavioral research, I aim to foster a more interdisciplinary approach. This synthesis broadens the toolkit available to researchers and paves the way for innovative insights and findings.

Complex Insights into Human Decision Making: At the heart of economics lies the study of human decision-making. With the tools and techniques at our disposal, I seek to shed light on the insights that arise from human economic-driven decisions and strategic interactions. Understanding these complexities is paramount, as it informs policies, strategies, and interventions that can have profound societal and corporate impacts.

Bridging Theory and Practice: As the world becomes more interconnected and dynamic, there is a need to ensure that economic theories are not just sound in principle but also applicable in real-world scenarios. I strive to bridge the gap between theoretical constructs and their practical implications by combining behavioral insights with experimental data and computational analysis.

In essence, this research is motivated by the desire to push the frontiers of economic understanding, leveraging the strengths of technology, methodology, and interdisciplinary collaboration.

1.2 Main Contributions to the Literature

Across the five selected publications composing this thesis, the goal was to advance the field of economics through behavioral and experimental lenses, both in terms of findings and innovative methods, focused on robustness, generalizability, and reproducibility. Therefore, each piece of work provided a unique contribution. The author has conceptualized and executed the research reported in each paper, being greatly supported by the contributions of their respective co-authors.

The paper **Analyzing the Impact of Strategic Behavior in an Evolutionary Learning Model Using a Genetic Algorithm** introduces a novel heuristic-based simulation model that integrates game theory and Genetic Algorithms to explore strategic behavior and economic learning. The research is driven by the observation that strategic scenarios often evolve over time, prompting the need for a model where games can adapt based on agents' behavior.

As heuristic methods are often built for specific purposes, the paper outlines the theoretical procedure for constructing the simulation model. The model is designed to analyze 144 unique 2×2 games and three distinct strategy selection rules: Nash equilibrium, Hurwicz rule, and a Random selection method. The primary aim is to understand how strategic behavior influences the transformation of dynamic decision-making scenarios. The analysis focuses on different decision rules' performance in this evolutionary learning process, highlighting the consistent transformation of games and behavioral traits.

Furthermore, the originality of this approach lies in an algorithm that modifies the games based on their overall outcomes rather than altering the strategies or player-specific traits. This unique perspective offers insights into the evolution of optimal outcomes in various choice scenarios. The results from the study underscore the significance of strategic behavior in evolutionary learning, with findings indicating optimal player scenarios for both the Nash equilibrium and Hurwicz rules. The paper also observes the frequent transformation of games, allowing agents to coordinate their strategies and achieve stable optimal equilibria. Another critical observation is the evolution of game populations into groups of fewer repeating isomorphic games with strong preceding game characteristics.

In essence, this paper contributes to the literature with a comprehensive exploration of the impact of strategic behavior in an evolutionary learning model, offering valuable insights into the dynamics of game theory and economic learning through a methodologically rigorous simulation model.

The subsequent two papers are derived from experiments undertaken within the same research initiative. Although they share a common origin, their focuses diverge significantly. The first paper adopts a game-theoretical perspective, analyzing equilibria through behavioral dynamics. Meanwhile, the second paper introduces the notion of dark personality traits, examining how game mechanics, individual attributes, and contextual factors influence decision-making. Detailed overviews of each paper will follow next.

Stationary Equilibria in Behavioral Game Theory: An Experimental Analysis of Inspection Games, investigates the dynamics of inspection games, a subset of non-cooperative game theory, where strategic interactions occur between two actor profiles: an inspector and the inspected. The study's primary focus is understanding how behavioral traits, such as loss aversion and inequity aversion, influence the equilibrium behavior in these games.

The research framework is designed to capture interactions influenced by psychological incentives, including moral connotations, punishment, and regret. This approach allows for a comprehensive exploration of behavioral parameters and their dynamics in strategic decision-making.

A significant aspect of this study is the introduction of both neutral ("context-free") and information-loaded ("in-context") frames. This design choice aims to measure the effects of information framing on perceptions and equilibria, especially in scenarios with potential ethical implications.

Methodologically, the paper evaluates several stationary equilibrium concepts: Nash Equilibrium, Quantal Response, Action Sampling, Payoff Sampling, and Impulse Balance. The analysis extends to understanding the models' characteristics and performance at different aggregation levels. Additionally, the paper introduces modified versions of these equilibria concepts, integrating behavioral traits to provide more flexible models and accurate predictions.

The results indicate enhanced predictive performance in cyclic games when integrating behavioral traits into stationary equilibrium models. While traditional models faced challenges, the modified versions aligned more closely with observed behavior. Notably, the inclusion of loss aversion as a model parameter in these modified concepts aligned with the actual behavioral tendencies of participants. As for framing, the study revealed minimal impact, with detailed information subtly nudging participants to opt for riskier strategies.

This paper extends the foundational literature on stationary equilibria by presenting modified concepts, examining the impact of behavioral traits and framing effects, and offering an in-depth assessment of model parameters and their evolving dynamics. The primary contribution of this research is a broader and more detailed perspective on equilibria in completely mixed games.

I further explore the inspector game framework in the subsequent manuscript, titled **Understanding Dark Personality Traits and Strategic Choices in an Inspection Game: Insights from Machine Learning and Causal Inference**. This paper provides a more comprehensive analysis of framing effects and introduces the concepts of dark personality traits and dishonesty. While it draws upon the repeated game dataset from the preceding paper, it extends the research by integrating a one-shot game dataset with a larger sample of participants.

The employed experiment utilizes the same design with information-loaded and neutral frames. Further, it employs the D-Factor and dishonesty questionnaires, which quantify dark personality traits, shiting the focus on moral and behavioral aspects. The objective is understanding how framing, game mechanics, and personal elements influence decisions and the interaction between experimental participants in repeated and one-shot settings.

The study's methodology began with a statistical evaluation of game behavior, followed by applying advanced machine learning models to analyze the influence of personality traits and framing effects. Random forest models were used to assess the impact of contextual and personal variables on decision-making. Additionally, causal forest was employed to provide a detailed analysis of framing, quantifying the causal treatment effects and examining how participants' traits influenced these effects.

The primary findings emphasize the dominant role of game mechanics in decision-making. The D-factor, representing dark personality traits, emerges as a secondary factor influencing decisions, especially in the inspector profile under framed conditions. This profile displays a trend towards higher retaliation when subjected to certain conditions.

This paper's main contributions include a novel analysis of dark personality traits within the context of inspection games, filling a gap in the existing literature. The research employs a robust methodology designed to deliver complex insights that might remain obscured with conventional methods. Furthermore, the study offers a holistic understanding of decision-making, factoring in the strategic interaction and the influence of contextual and personal elements. For practical implications, the paper sheds light on behaviors related to cheating and punishment. Such insights can be valuable for policy formulation, especially in areas that require behavioral regulation, and can also be instrumental in designing interventions to mitigate dishonest behaviors.

The following paper, titled **Trust in the Machine: Contextual Factors and Personality Traits Shaping Algorithm Aversion and Collaboration**, explores the intricate relationship between contextual factors, personal variables, and the phenomenon of algorithm aversion in decision-making. The study is set in an experimental environment where subjects are presented with the option to delegate decisions to a Reinforcement Learning algorithm while navigating a multi-armed bandit problem. The experiment is designed to discern the superior option with hidden expected values.

A standout feature of this research is its distinctive experimental design, allowing participants to interact directly with individual instances of Reinforcement Learning models. This design is further enriched with four treatments: baseline, explanation, payment, and automation. Each treatment offers insights into different facets of human-algorithm interaction. For instance, the explanation treatment provides a non-technical algorithm description to enhance transparency and reduce the "black box" perception. The payment treatment introduces a cost associated with algorithmic support, exploring the psychological implications of financial incentives in decision delegation. The automation treatment, on the other hand, offers participants the convenience of continuous algorithmic selection, reducing the task effort.

Methodologically, the paper employs a combination of statistical, regression, and machine learning techniques, including Logistic Regressions, Random Forest, Gradient Boosting Machines, and Uplift Random Forest classifiers. The analysis evaluates the influence of personal and contextual factors on participants' decisions. Subsequently, the study employs a causal inference approach to disentangle the effects of context from individual characteristics. These tools examine the nuanced nature of decision delegation behavior, revealing robust insights such as the negative impact of payment on delegation and the positive influence of full automation.

On the personal dimension, the study incorporates a range of psychological and demographic measures, including the Big Five Personality traits, Locus of Control, and Generalized Trust. Findings indicate that age, extraversion, openness, neuroticism, and locus of control are pivotal in shaping delegation decisions. Notably, female participants exhibited a heightened sensitivity to algorithmic errors.

Furthermore, the paper unveils several novel insights. The detrimental effect of even a small payment on decision delegation suggests that the perceived cost can outweigh the perceived benefits of algorithmic support. Additionally, gender effects emerge as a significant factor, with female participants reacting more strongly to algorithmic mistakes. The findings offer valuable insights into crafting user-centric AI designs that foster collaboration while minimizing aversion.

Lastly, the paper titled **Motivated Sampling of Information: Experimental Data and Agent-Based Modeling in a Bayesian Framework** explores the idea of information sampling, where objective and subjective criteria exist. The paper introduces an experimental framework that addresses the phenomenon of motivated reasoning, termed motivated sampling. Participants in the study are presented with a binary sampling and decision task, where they must discern information from two "computers" generating numbers from distinct distributions. The objective is to identify the "high distribution" computer. The experiment introduces externalities to influence participants' decisions, thereby inducing subjective preferences. Moreover, the type of feedback provided to participants varies, offering an understanding of its influence on decision-making.

The main findings highlight that female subjects, for instance, sample more intensively in scenarios with negative externalities or when provided Bayesian posterior feedback. Additionally, subjects exhibit a pronounced margin of motivated sampling, especially when they perceive the option with a positive externality as correct. This behavior underscores a preference to sample from options that align with objective and subjective criteria.

Furthermore, the study employs a simulation model that mimics the task of sampling for information. This model incorporates agents based on Attraction-Based Reinforcement Learning, ϵ-greedy strategies, and the Upper Confidence Bound method. These agents, in essence, learn the sampling process to optimize rewards from subsequent computer selections. A comparison with a random sampling agent further enriches the analysis.

In essence, this paper offers a novel perspective on information sampling strategies, unraveling the mechanisms of motivated sampling. It comprehensively explains the overlap between information sampling and motivated reasoning, emphasizing the significance of subjective preferences, feedback types, and gender differences in decision-making scenarios. The findings contribute to the broader theme of motivated reasoning and highlight its specific application in situations where information sampling is relevant.

In summary, the papers presented in this dissertation integrate across three foundational pillars: the experimental, the behavioral, and the computational domains. These domains, while diverse, provide a holistic perspective. Table 1.1 briefly summarizes each paper's alignment with these pillars.

The employed techniques covered laboratory, field, and computer-simulated experiments within the experimental facet. The behavioral dimension spans a broad spectrum, from exploring classical strategic decision-making behavior to assessing well-established behavioral economics concepts like loss aversion and fairness. This dimension further touches upon the dark facets of personality, the big five personality traits, elements of trust, control, and underlying motivations behind information acquisition and decisions. The computational domain encompasses a spectrum of artificial intelligence methodologies from genetic algorithms to optimization processes, an array of machine learning techniques including ensemble learning using random forests, gradient boosting, reinforcement learning strategies, and diving into causal machine learning with techniques like causal forests and uplift modeling.

1.3 Structure of this Document

This thesis is structured to present the five key papers discussed sequentially in the following chapters. Each paper comprises its abstract, introduction, methods, theoretical foundations, experiment designs, and pertinent conclusions and discussions. Each paper distinctly articulates its aims, employed methods, and derived results. Subsequent to each paper, an appendix offers further insights into the research. The literature used across the papers is consolidated and presented at the end of the document. The ensuing chapter, 7, encapsulates the work's primary takeaways, contributions, and conclusions.

Paper 1 - Simulating Economic Learning in Dynamic Strategic Scenarios with a Genetic Algorithm (chapter 2): *Experimental Approach:* Simulated experiments using computer agents. *Behavioral Approach:* Analysis of simulated strategic behavior in 144 2×2 games. *Computational Approach:* Specially designed genetic algorithm to modify games where computational agents interact based on optimal decision behavior. *Co-author:* Thomas Pitz *Status:* Published in Computational Economics (DOI: 10.1007/s10614-022-10348-1)
Paper 2 - Stationary Equilibria in Behavioral Game Theory: An Experimental Analysis of Inspection Games (chapter 3): *Experimental Approach:* Laboratory experiment, between-subject design with two treatments employing information-loaded and neutral frames. *Behavioral Approach:* Analysis of Stationary Equilibria concepts with traits like loss aversion and inequity aversion in completely mixed cyclic games. *Computational Approach:* Development of new equilibria concepts formulated as optimization models that include behavioral parameters. *Co-authors:* Thomas Pitz, Wolf Gardian, Deniz Kayar & Jörn Sickmann *Status:* Submitted for publication in November 2023
Paper 3 - Understanding Dark Personality Traits and Strategic Choices in an Inspection Game: Insights from Machine Learning and Causal Inference (chapter 4): *Experimental Approach:* Laboratory and online experiments, between-subject design with two treatments employing information-loaded and neutral frames. *Behavioral Approach:* Analysis of strategic and moral decision-making linked with dark personality traits. *Computational Approach:* Analysis using Random Forests and Causal Forests with Double Machine Learning estimators for causal inference. *Co-authors:* Leon Houf, Thomas Pitz & Christiane Schwieren *Status:* Submitted for publication in November 2023
Paper 4 - Trust in the Machine: Contextual Factors and Personality Traits Shaping Algorithm Aversion and Collaboration (chapter 5): *Experimental Approach:* Online experiment, between-subject design with four treatments. *Behavioral Approach:* Examination of the algorithm aversion phenomenon via contextual factors and personal insights. *Computational Approach:* Analysis using Random Forests, Gradient Boosting Machines, and Uplift Random Forest models for causal inference. *Co-authors:* Leon Houf, Thomas Pitz, Christiane Schwieren & Jörn Sickmann *Status:* Submitted for publication in September 2023
Paper 5 - Motivated Sampling of Information: Experimental Data and Agent-Based Modeling in a Bayesian Framework (chapter 6): *Experimental Approach:* Online experiment, between-subject design with a 3×3 treatments split with varied feedback and externality conditions. *Behavioral Approach:* Analysis of motivated decision-making and Bayesian behavior. *Computational Approach:* Agent-based models using Reinforcement Learning techniques: Attraction-Based RL, ϵ-greedy strategies, and Upper Confidence Bound. *Co-author:* Leon Houf *Status:* Will be submitted for publication

Table 1.1: Summary of Papers and Approaches

2. Simulating Economic Learning in Dynamic Strategic Scenarios with a Genetic Algorithm

Authors

Vinícius Ferraz & Thomas Pitz [1]

Abstract

This study presents an experimental approach to strategic behavior and economic learning by integrating game theory and Genetic Algorithms in a novel heuristic-based simulation model. Inspired by strategic scenarios that change over time, we propose a model where games can change based on agents' behavior. The goal is to document the model design and examine how strategic behavior impacts the evolution of optimal outcomes in various choice scenarios. For diversity, 144 unique 2×2 games and three different strategy selection criteria were used: Nash equilibrium, Hurwicz rule, and a random selection technique. The originality of this study is that the introduced evolutionary algorithm changes the games based on their overall outcome rather than changing the strategies or player-specific traits. The results indicated optimal player scenarios for both The Nash equilibrium and Hurwicz rules, the first being the best-performing strategy. The random selection method failed to converge to optimal values in most of the selected populations, acting as a control feature and reinforcing the need for strategic behavior in evolutionary learning. Two further observations were recorded. First, games were frequently transformed so agents could coordinate their strategies to create stable, optimal equilibria. Second, we observed the evolution of game populations into groups of fewer (repeating) isomorphic games with strong preceding game characteristics.

Keywords

Game Theory, Simulation, Genetic Algorithms, Economic Learning, Artificial Intelligence

2.1 Introduction

Our world is constantly going through systemic transformations. Technological and scientific advancements, in combination with changes in economic, political, sociological, and other factors, result in new decision contexts in which strategic interaction occurs. By analyzing such situations, one can observe the ever-arising need for individuals to update their beliefs and adapt to the new informational and strategic structures. As an example, Freedman (1998, 2017) reported the example of socioeconomic and political transformations occasioned by the development of technology and communication structures, resulting

[1] **Status:** This paper was published in Computational Economics, in December 2022, under the DOI: `https://doi.org/10.1007/s10614-022-10348-1`.

in changes in how nations engage in warfare conflict. When making decisions, economic agents will act according to their interests, as well as the actions of other agents and information available in the given scenarios, as outlined by Von Neumann and Morgenstern (1953).

As stated by Axelrod (1997), one can understand the properties of complex social and economic systems by applying simulations. The nature of human interaction is often modeled and analyzed in computational models of society, which introduce autonomous agents that interact with one another and the environment into which they are placed according to predefined rules (Billari, Fent, Prskawetz, & Scheffran, 2006). Adding dynamics to models of strategic interaction, social learning, and the evolutionary process are often simulated by introducing evolutionary biology concepts, as outlined by Gintis et al. (2000). This evolutionary approach introduces the notion of predefined strategies that are repeatedly applied in an evolutionary process, operating dynamically on the distribution of behavior (Weibull, 1997).

Game theory models describe strategic scenarios and behavior (Von Neumann & Morgenstern, 1953). In game theory, political or socioeconomic conflicts or crises are often modeled in a strategic form by matrix games or in an extensive form by game trees. One often restricts oneself to a fixed game that does not change significantly over time. Prominent examples are the analysis of the Cold War as a Prisoner Dilemma (Plous, 1993) or the Cuba crisis as a Chicken Game (B. Russell, 1959). In reality, however, it is observable that the strategic character of conflicts or crises changes over time. A crisis modeled as a Prisoner Dilemma can intensify into a Chicken Game with higher conflict potential or transform into a less conflictual Stag Hunt (Skyrms, 2004) or Harmony Game (Bruns, 2010). Therefore, it would be appropriate to describe these strategic changes by transforming the original game into a new game. In empirical settings, fundamental behavioral changes are observed when making decisions that can affect not only the strategic behavior of the agents involved but also environmental conditions and individual preferences. Heckathorn (1996) documented the transformation of games with dynamic interaction, where changes in decision-influencing factors changed the whole structure of the initial game. Similarly, Simpson (2004) empirically demonstrated how behavioral factors, such as social preferences, can transform one game into another (see also Hayashi, Ostrom, Walker, and Yamagishi (1999); Kollock (1998)).

Motivated by this fact, in the present paper, we introduce a novel heuristic procedure that describes these changes in strategic interaction scenarios with Genetic Algorithms, a search procedure inspired by the process of biological evolution (D. E. Goldberg & Holland, 1988; Holland, 1992). Genetic Algorithms have been employed in economics-based problems since their introduction and are regarded as a powerful tool for finding optimal solutions over complex search spaces (Schmertmann, 1996), as well as a method that shows remarkable results for the simulation of decision behavior that is in line with empirical observations in similar research frameworks (Arifovic & Ledyard, 2012; Lensberg & Schenk-Hoppé, 2021; Manson, 2006). Given a pool of 2×2 strategic-form games, the games, represented as binary sequences, are transformed by a Genetic Algorithm depending on the players' experience with the games in the pool. In each round, after making decisions, the game is evaluated based on the players' payoffs. This process determines the probability of staying in the pool, or being replaced by another new game created via crossover and mutation. We describe the dynamics of the game pools related to the different types of strategy selection rules adopted by the players, establishing the notion of strategic behavior for the agents. The populations are defined based on two rules for aggregating sets of games. The first rule is based on topological proximity, using the periodic table of 2×2 games concept introduced by Robinson and Goforth (2005); the second rule clusters games by similar characteristics, based on the families categorization introduced in Bruns (2010, 2015a).

The introduced model focuses on analyzing the impact of strategic behavior in this evolutionary learning process, where games are allowed to change over time and the performance of different decision rules. We aimed to document the implementation, testing, and assessment of different decision-making rules and their influence on dynamic game populations. For a comprehensive analysis, the simulation model described here analyzes 144 unique types of 2×2 games and three distinct strategy selection rules: Nash equilibrium, Hurwicz rule, and a Random selection method. The goal is to outline how strategic behavior affects the transformation of dynamic decision-making scenarios by pairing different strategy selection rules with distinct populations of 2×2 games. The analysis of the simulation results focused on convergence speed (optimal utility levels reached) using different combinations as a performance measure, as the encoding of strategies and convergence process are seemingly interconnected (Dawid & Kopel, 1998). We have also documented the findings derived from the simulation process, including the consistent transformation of games and behavioral traits highlighted during the process [2].

2.2 Evolutionary Computation Overview

Evolutionary computation methods arose from taking inspiration from biological mechanisms to design and implement computer-based problem-solving systems (Spears, Jong, Bäck, Fogel, & Garis, 1993). This collection of methods allowed the creation of evolving and adaptive solutions to complex problems, especially the ones that impose challenges to traditional algorithms, such as randomness, chaotic disturbances, and complex non-linear dynamics, as outlined by Fogel (2000). The family of evolutionary algorithms contains several different methods, each with its particularities, but all of them share a connection with biological evolution. Among the most known methods in the literature are Genetic Algorithms (Holland, 1992), Genetic Programming (Koza et al., 1992), Differential Evolution (Storn & Price, 1997), Evolution Strategies (Rechenberg, 1978) and Evolutionary Programming (Fogel, 1998). For an overview of each of these methods, see Slowik and Kwasnicka (2020). The literature discussion that follows next outlines applications of some of these methods in similar contexts, highlighting the elements that led to adopting a Genetic Algorithm as an appropriate method to simulate the dynamic transformations of games over time.

[2] The documented source code, data, and resources used in this paper are published in the CoMSES Library. The material is available to download at: `https://doi.org/10.25937/smg0-0t92`.

2.2.1 Evolutionary Models: Applications for Learning, Strategic Interaction, and Optimization

As C. Camerer (2003) outlined, some aspects of learning are sometimes overlooked by economic theory. If perfect information and rationality are assumed, the equilibrium point will always be known from the beginning, and people will only modify the equilibrium if information changes. Moreover, C. Camerer and Weigelt (1988) emphasized the importance of achieving better outcomes in experimental games, especially when dealing with scenarios having potentially inefficient equilibrium outcomes, such as trust games, public goods games, beauty contests, and others. Consequently, well-formulated economic learning theories are crucial in providing predictive power, coherence, and concomitantly revealing new insights (C. Camerer, 2011).

The multidisciplinary combination of game theory and genetic programming has grown in several distinct fields, from economics and sociology to computer science and natural sciences such as biology. Evolutionary game dynamics provide comprehensive frameworks for studying interaction, learning, and evolution (Roca, Cuesta, & Sanchez, 2009). In addition, in contrast with the neoclassic assumption of perfect rationality, economic models of learning provide the possibility to study agents as they learn and update their beliefs since the application of an evolutionary model assumes that strategies can change over time (Baddeley, 2018). According to Axelrod et al. (1987), individuals cannot thoroughly analyze the situation and calculate optimal strategies when interacting in complex environments. Alternatively, strategies are updated and based on achieved results, highlighting how a Genetic Algorithm can be particularly adept as a learning mechanism for creating effective strategies. The given approach serves as an inspiration for the analysis performed in this article. The following are some of the relevant constructs.

Holland and Miller (1991) stresses that the employment of artificial adaptive agents in economic theory can help us understand real-world economic issues by enabling the free exploration of system dynamics under controlled conditions and the opportunity to check several unfolding behavioral patterns. Furthermore, Dawid (1999) argued that the decentralized structure of Genetic Algorithms, which naturally resembles a group of interacting economic agents, is well-suited to simulate the behavior of economic systems.

Isaac (2008) provided an introductory overview of an agent-based model using Genetic Algorithms in the iterated Prisoner's Dilemma, reporting variations in the payoff structures that create new player types, introducing an interaction between payoff cardinality and players' attributes. Chmura, Kaiser, and Pitz (2007); Pitz, Chmura, et al. (2005) presented a novel simulation model for analyzing action patterns in social systems mainly based on the concepts of Genetic Algorithms and the Theory of Action Trees (Goldman, 1971). They explained how the emergence and disappearance of actions could be described with a uniform algorithm, succeeding in endogenously eliciting comprehensive changes in the agents' behavior. Manson (2006) documented experiments exploring the concept of bounded rationality, stating that Genetic Algorithms are an appropriate tool to model actors that are not perfectly rational, that is, addressing characteristics of human decision-making such as cognitive limits, learning, and innovation.

Similarly, the algorithm for optimization problems reported by Yang (2017) build on a similar conceptual framework. Their main idea is a game theory-inspired evolutionary model that updates the strategy sets by replacing individuals of the population with better-performing offspring generated by replication or belief learning operators, creating a model that outperforms four other algorithms often used for similar purposes. Pereira et al. (2020) also introduced a constrained optimization model that explored two ideas, the first being a Genetic Algorithm with social interactions (for diversification of solutions in the selection process). The second model consisted of game-based crossovers (tournament simulations for more diverse offspring). The presented construct demonstrated robust performance when compared to traditional methods in the engineering design optimization process. Continuing on the topic of optimization.

Savin and Egbetokun (2016) formulated an agent-based model of innovation networks with endogenic absorption capacity, where dynamic cooperation for knowledge can occur between different agents, represented as firms, with different knowledge positions. In their simulation model, the authors applied a Differential Evolution algorithm to find optimal investment budget decisions regarding trade-offs between cooperative and non-cooperative scenarios. Their findings demonstrate that networks generated with the model display small-world properties, which tend to be efficient structures for knowledge distribution. Another interesting observation is that firms with higher absorptive capacity tend to be better positioned within their networks, ultimately demonstrating that their ability to learn drives network performance effects.

With a similar objective to this paper, Savin, Blueschke, and Blueschke-Nikolaeva (2018) introduced a meta-heuristic approach for solving non-linear dynamic games, proposing a method that allows the analysis of more realistic strategic scenarios. The method can solve the standard version of the introduced game, like other traditional techniques, and solve non-standard extensions of the problem (inequality constraint and asymmetry in penalties) by identifying optimal equilibrium strategies, both cooperative and non-cooperative. The proposed procedure combines Differential Evolution (for individual optimal strategies) and Approximation of a Nash Equilibrium. The set consists of a three-player macroeconomic game between two groups of countries exercising fiscal policy and one joint central bank. The results shed light on a more realistic analysis of strategic scenarios for policy insights and finding optimal strategies, contrasting with traditional methods. In further developments on the Differential Evolution method application, Savin and Blueschke (2016) introduced a model to solve optimal control problems, addressing the limitations of classical methods in specific situations. The model's performance demonstrated a superior optimization of expected outcomes, strengthening the claim that heuristic methods are well-suited to navigate complex search spaces and find good approximations to global optima. Blueschke, Savin, and Blueschke-Nikolaeva (2020) later extended on this subject by introducing a novel Differential Evolution-based method for solving optimal control problems with passive learning. This learning method models the observation of the world's current state and employs new information to improve the system's general knowledge. The proposed approach does not imply the modification of the original problems and provides more robust results regarding the learning process.

Bullard and Duffy (1998) documented a macroeconomics experiment using a Genetic Algorithms-based learning model to simulate the behavior, outlining that a population of artificial agents can coordinate depending on the information structures they are inserted into, as chaotic and complex structures tend to hinder coordination. Another related framework was developed by Gooding (2014), who formulated a simulation model for capturing evolutionary trends observed in society, such as wealth aggregation, inequality, and climate change. Where experimental data verify changes in actions, surroundings, and decision-making, such social trends remain resilient and difficult to alter, according to the study, offering insights into how to impact social development. Macedo, Godinho, and Alves (2020) applied a Genetic Algorithm to optimize trading strategies, which outperformed the analyzed market indicators by employing a more comprehensive search space than traditional methods.

In a similar context, Glynatsi, Knight, and Lee (2018) used an evolutionary game theoretic model in the ecology field to examine the interaction between poachers and wildlife. The model analysis reported how the devaluation of rhino horns would likely lead to higher poaching activity and that such an approach was only practical when combined with disincentives, intending to contribute to informing debates on the issue with scientific facts.

Arifovic and Ledyard (2011) introduced an evolutionary learning model with relatively good performance at matching the behavior of agents engaged in repeated strategic interactions when the behavior converges to a Nash equilibrium state. The authors state that most games do not require sophisticated strategies, except for the case of coordination games, which reduces the model's performance. Arifovic and Ledyard (2012) later reinforced the predictive power of evolutionary learning methods by introducing a comprehensive model able to generate data quantitatively similar to the empirical values, focused on the contribution mechanism of a public goods game. Price (1997) also reported a good performance from Genetic Algorithms in searching for equilibria in standard games from industrial organization theory, such as Bertrand and Cournot competition scenarios.

Koza (1994); Koza et al. (1992) provided an early framework for Genetic Programming that introduced the notions of learning by modeling agents and their learning behavior over time. The author defines adaptive learning as the process of changing the structure of a potential solution. Hence, it performs better in its environment, where positive changes are rewarded and negative changes are discouraged by the underlying fitness function. Similarly, S.-H. Chen, Duffy, and Yeh (2005) introduced a comprehensive game theory and Genetic Algorithms framework that approach several topics present in this research, such as coordination, adaptive learning, and equilibrium selection. The authors compared the behavior of computational agents to human subjects. They concluded that the behavior was remarkably similar in the applied experiment, supporting the idea of Genetic Algorithms as a credible tool to model human behavior.

Lensberg and Schenk-Hoppé (2021) studied the process of learning in one-shot multiple

2×2 games, where the agents never only see each game once and should learn to find optimal strategies based on information acquired across games. The author proposes a solution concept based on multiple artificial agents that learned how to play the games through Genetic Algorithms. The proposed theoretical model is reported to perform well, in line with intuition and empirical evidence.

Other interesting applications of the combination of Genetic Algorithms and game theory described in published literature include practitioners in other distinct and diverse fields, such as engineering (Périaux, Chen, Mantel, Sefrioui, & Sui, 2001), energy (Castillo & Dorao, 2013, 2012; Mohamed & Koivo, 2011), communications (Kusyk, Sahin, Uyar, Urrea, & Gundry, 2011), land usage (Liu et al., 2015), biology (Hamblin & Hurd, 2007) and ecology (Hamblin, 2013).

In summary, the literature suggests that Genetic Algorithms are an appropriate model for adaptive learning and optimizing strategic decisions. It performs well in problems of strategic interaction models (such as ours) while incorporating behavioral traits that are close to empirical findings in experiments with human behavior. Genetic algorithms suit our objective since they allow us to manipulate binary sequences under imposed constraints. In this case, one can transform the numerical structure of game elements so that essential characteristics are taken into account, as well as the outcomes of the decisions performed by the agents, expressed by the fitness function generated by different strategy selection rules. In this way, repeated decisions of the agents can influence the transformation of the games to directions that are consistent with the agents' behavior, providing us the necessary building blocks to simulate a situation where games can transform and analyze how the behavior of the agents influences these transformations.

2.3 Game-Theoretical Features

This article adopted the standard representation of strategic-form games as a model of simultaneous interaction between two agents, denoted by a 2×2 matrix. This form encompasses the following elements: the (two) players, who are the parties making the decisions; the strategies that can be selected by each player (two for each) and the payoffs being the rewards received as a function of the chosen strategy (Robinson & Goforth, 2005; Von Neumann & Morgenstern, 1953). The strategic representation of games focuses on static analyses while overlooking dynamic aspects such as the order of the players' moves, changes in the moves, and the informational structure. This approach suggests the strategies that are more likely to be used by each player or alternatively recommend to players which strategies to choose in similar scenarios (Maschler, Solan, & Zamir, 2013). In Robinson and Goforth (2005) 's notation, the players in the context of this article are named after how the strategy profiles are organized, with the player ROW's strategies displayed in the rows of the matrix and player COL's strategies in the columns, respectively.

The adopted classes of 2×2 strategic-form games were based on the "periodic table" categorization provided in Robinson and Goforth (2005), which formally connects all ordinal rank games with distinct player preferences since swaps topologically linked the games in adjoining payoffs. The space of 2×2 is infinite, though as we were only interested in ordinal preference relations, we can concentrate on classes of isomorphic games, where for each class, we can choose one representation of the form $\{1, 2, 3, 4\} \times \{1, 2, 3, 4\}$. Respectively, as Robinson and Goforth (2005) have demonstrated, there are 576 ways to arrange two sequences of four numbers in a bi-matrix scheme. The ordinal structure of a game does not change by switching rows, columns, or both simultaneously; the 576 games can be reduced by a factor of 4 to 144. For a broader representation of strategic scenarios, all 144 unique classes of games - including a wide range of well-known applied game theory situations such as the Prisoner's Dilemma, Chicken game, Stag Hunt, Battle of Sexes, and several others - are included in the simulation model.

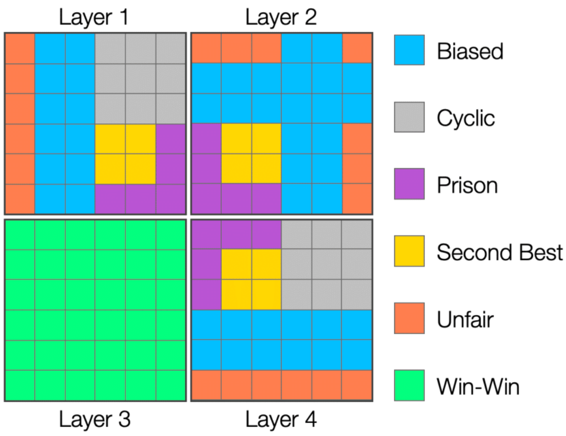


Figure 2.1: Representation of the Game Families and Layers in the Periodic Table of 2×2 games (Bruns, 2010, 2011, 2015a, 2015b)

Game Family	Nash Equilibria	Pareto-optimal	Dominant Strategies	Count	%	Details
Biased	1	2	0-2	44	31%	One player gets the best and the other second best outcome
Cyclic	0	2-4	0	18	13%	In each cell one player would prefer to change their move, no pure strategy equilibrium
Prison	1	2-3	1-2	15	10%	Dominant strategies based on individual incentives leads to a worse outcome than cooperation
Second Best	1	3	1-2	12	8%	Both players gets the second best outcome
Unfair	1	2-3	1-2	19	13%	One player gets the best and the other the second-worst outcome
Win-Win	1	1	1-2	36	25%	Both players get the best outcome

Table 2.1: Characteristics of games categorized in families (Bruns, 2015a, 2015b)

In complement to the periodic table approach, Bruns (2010) further categorized the games by similarity in Layers, which outlines topological proximity, and Families, which groups similar games (based on equilibria and payoff structures, see Table 2.1). The games were split into populations following both grouping rules, as the scheme described in Figure 2.1 and Table 2.1.

2.4 Simulation Model Design

To have dynamic and transforming games, this simulation model processes the game-theoretical elements described earlier by allowing the games to change as a function of the players' strategic behavior. The Genetic Algorithms method becomes a fundamental building block, an enabler of these dynamic transformations. Genetic Algorithms are a class of search heuristic programs inspired by the process of natural evolution (Holland, 1992; Holland & Miller, 1991), being part of a broader set of comparable methods named Optimization Heuristics. The latter term, usually linked to algorithms inspired by nature, is defined by Gilli and Winker (2009) as methods that provide high-quality approximations to the global optimum, robust to changes, not too sensitive to parameters, easily deployable to many types of problems and might be stochastic, but without subjective elements. In Genetic Algorithms, the search space for potential solutions imitates the process of biological evolution. There are many variants of methods that are considered Genetic Algorithms. However, usually, this class of algorithms shares the following characteristics: having a population of individuals as potential solutions represented as binary strings, an objective function (fitness or cost), and the three types of genetic operators: selection, crossover, and mutation (Holland, 1992; Leszek, 2008; Mitchell, 1998; Slowik & Kwasnicka, 2020). The following sections will explain these individual elements in detail as we describe our implementation method.

Our approach allows the agents to modify their environment, especially regarding how they decide, to measure the progress of the genetic learning process in terms of strategy performance. The simulation process basically consists of playing the games in the designated populations and adopting one of three different strategy selection rules: (1) Nash equilibrium, (2) Hurwicz rule, and (3) random selection. After the strategies' selection, the games were assigned fitness scores based on the aggregated payoffs from both agents' choices; therefore, the games can be selected for replication in a way that favors higher-performing combinations of strategies. These preferences were implicitly expressed by the fitness function, which defined the quality of the selected strategies in terms of gained utility. In the next step, the games were processed by the Genetic Algorithm, which selects two games (parents) from the pool and generates a new game via the crossover and mutation operators. The new games were inserted back into the population, and the process was iterated a fixed number of times.

To manage the complexity and the processing requirements, the population numbers were kept constant throughout the iterations of the evolutionary model, which means when a new game is added, another game is excluded.

2.4.1 Populations and Data Structures

The encoding of games is based on a vector representation (payoff structure reproduced by integer vectors, as in Figure 2.3). The analysis was rooted in two primary divisions of game pools based on families and layers. On the one hand, the first restriction derived from Robinson and Goforth (2005) 's division of layers in their definition of the periodic table of 2×2 games, which considers all the 36 neighboring games according to the number of payoff swaps (see Figure 2.1). The entire space (all 144 games, not allowing ties) was also handled as one distinct population, being processed apart from the others, allowing

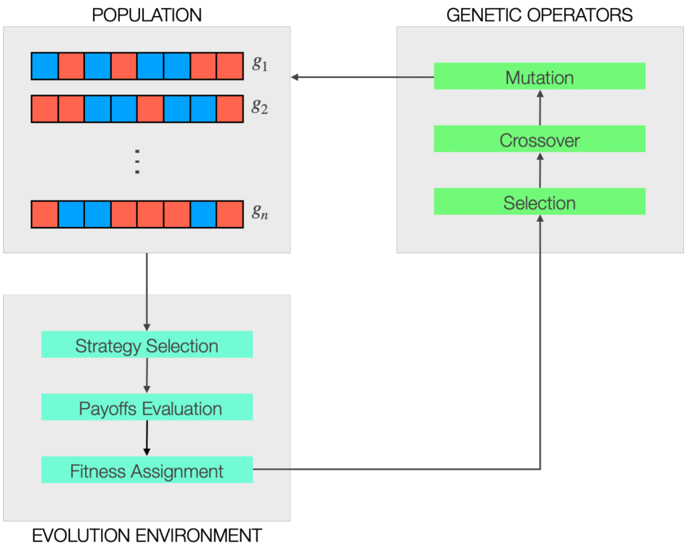


Figure 2.2: Simulation Model Overview

us to analyze the population's results using the complete set of available characteristics in the pool.

On the other hand, the second restriction was based on Bruns (2015a) 's game families' categorization, split into six groups (as in table 2.1). These two restrictions separated our groups of games into eleven distinct populations, processed and analyzed individually.

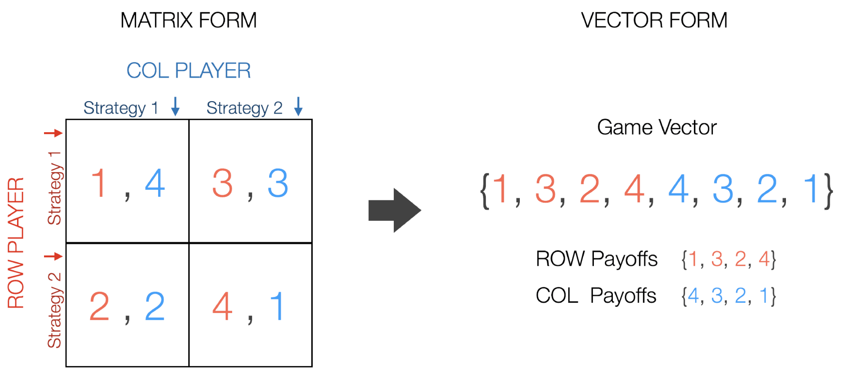


Figure 2.3: Game Vector Encoding Scheme

2.4.2 Binary Encoding

We adopted a binary representation of the vector form for the genetic operations. As our payoff matrices were represented by integers between1 and 4, each possible payoff value has been encoded with one pair of bits. Consequently, the binary encoding was based on four possible sequences using a double binary representation:

$$
\begin{aligned}
0,0 &\rightarrow 1 \\
0,1 &\rightarrow 2 \\
1,0 &\rightarrow 3 \\
1,1 &\rightarrow 4
\end{aligned} \tag{2.1}
$$

As an example of a vector transformed to binary, the representation of the Prisoners' Dilemma game (game in Figure 2.3) is defined as:

$$[1,3,2,4,4,3,2,1] \rightarrow [0,0,1,0,0,1,1,1,1,1,1,0,0,1,0,0] \tag{2.2}$$

This encoding scheme was applied in all 144 games. All game binary strings contained 16 bits, with each pair of bits representing one of the payoff values in the matrix. The game vectors could only be modified in accordance with this framework; thus, the structural scheme remained unchanged.

2.4.3 Evolution Environment: Strategy Selection and Fitness

This section explains the processes that govern the agents' strategy selection rules and how the players should act in the scenarios presented. The three approaches for selecting strategies were used to measure the impact of distinct types of strategic behavior.

Nash Equilibrium Strategy Selector

The Nash equilibrium is a strategy profile in which each strategy is an optimum response to other players' strategies. This logic holds to pure and mixed-strategy profiles (probability distribution over the available choices). As expected, utilities are linear in terms of probabilities. If a player in Nash equilibrium utilizes a non-degenerate mixed strategy, it must be indifferent from all other pure strategies assigned with a positive probability. A strict Nash equilibrium exists when each player has a unique best response to its rivals' strategies (Fudenberg & Tirole, 1991; Maschler et al., 2013; Von Neumann & Morgenstern, 1953). In our games, we will encounter all Nash Equilibrium situations, with strict and weak equilibria and pure and mixed strategies.

Because there can be multiple equilibria in bi-matrix games, the approach selected is the support enumeration method (Knight & Campbell, 2018; von Stengel, 2007; Widger & Grosu, 2008). The support enumeration computes all equilibria for a degenerate 2×2 game $(A, B) \in \mathbb{R}^{m \times n^2}$, for all $1 \leq k_1 \leq m$ and $1 \leq k_2 \leq n$ (enumeration of all possible equilibrium strategies); for all pairs of support, (I, J) with $|I| = k_1$ and $|J| = k_2$. In other words, the goal was to find support for strategies played with a non-zero probability. At this point, the algorithm evaluated the best response condition, ensuring no better utility residing outside of the supports.

The steps outlined in equations (2.3) and (2.4) iterate through all potential support pairs.

$$\begin{aligned} \sum_{i \in I} \sigma_{ri} B_{ij} = v \text{ for all } j \in J, \\ \sum_{j \in J} A_{ij} \sigma_{cj} = u \text{ for all } i \in I. \end{aligned} \tag{2.3}$$

Sequentially, for considering mixed strategies:

$$\begin{aligned} \sum_{i=1}^{m} \sigma_{ri} = 1 \text{ and } \sigma_{ri} \geq 0 \text{ for all } i, \\ \sum_{i=1}^{n} \sigma_{ci} = 1 \text{ and } \sigma_{cj} \geq 0 \text{ for all } j. \end{aligned} \tag{2.4}$$

There were two equilibrium-selection rules for games with more than one Nash equilibrium. The first rile rule is based on the payoff dominance concept for simulating self-maximizing behavior (Harsanyi, Selten, et al., 1988). The equilibrium points might be finite or infinite, with at least one equilibrium (pure or mixed) for each game. Our structure identified three main configurations: pure-strict, pure-weak (when a stronger pure-strategy equilibrium was available), or mixed. The Nash equilibrium is payoff-dominant and henceforth selected if it is Pareto superior to all other possible equilibrium situations in a game. The second rule applied a simple random selection from the sample of computed Nash equilibria for each game. The objective was to enable the comparison of the maximization versus randomization approaches and measure the effects in the evolutionary process generated by more diverse procedures for Equilibrium selection.

Hurwicz Rule Strategy Selector

This study's second strategy selection rule employed the Hurwicz criterion (Hurwicz, 1951). This rule introduces a coefficient of realism, α, which serves as a tool for balancing pessimism and optimism in decision-making under uncertain scenarios, allowing decision-makers to account for different possible outcomes. The pessimistic option employs the *maximin* criterion, while the optimistic option employs the *maximax* criterion. The α parameter introduces a weighting factor between both extremes, simulating different degrees in behavior profiles.

Following Gaspars-Wieloch (2014) 's implementation, we applied the following formula.

$$h_j = \alpha \cdot w_j + (1 - \alpha) \cdot m_j, \tag{2.5}$$

resulting in the Hurwicz criterion, h_j, with α as the coefficient of realism, being $\alpha \in [0, 1]$. In this paper, 0 represents the pessimistic extreme, the risk-averse behavior, while 1 represents the optimistic extreme, or the risk-prone behavior (Colman, 2016). The optimal alternative between the two is expressed by:

$$h_j = \max_j \{h_j\} = \max_j \{\alpha \cdot w_j + (1 - \alpha) \cdot m_j\}. \tag{2.6}$$

This strategy selector introduced another model of behavior profiles and enriched the dataset to analyze simulation results. Three variations of the Hurwicz coefficient (α) were applied, simulating three distinct decision-making profiles: pessimistic (0.0), neutral (0.5), and optimistic (1.0). For other applications of the Hurwicz criterion in decision-making under uncertain scenarios, see Jaffray, Jeleva, Gains, and Paris (2007), Pažek and Rozman (2009), and Puerto, Mármol, Monroy, and Fernández (2000).

Random Strategy Selector

The third strategy selection mechanism consisted of a random choice of strategies for both players, simulating the total absence of strategic behavior. This method was added mainly as a control scheme, so one could assess if the populations were able to progress in terms of utility by not having any simulated decision rule and only relying on the maximization mechanism of the Genetic Algorithm - through selection, crossover, and mutation. The random selection was essential to outline the effects of strategic behavior introduced by the other two rules.

Payoffs and Fitness

Our fitness function reflected the players' preferences defined by the strategy selection rules applied at the game-playing stage, consequently taking the aggregated payoffs from each player as the overall game utility. The previously defined strategy selectors returned an array of probabilities (pure or mixed) of the agents selecting between the various available strategies. For systematic computations of payoffs, this algorithm employs a matrix multiplication method, using the dot product algebraic operation (Tanimoto, 2015). In this case, the strategies adopted by an agent during the execution of a game yielded probabilities distributions over the two possible strategies, expressed as the state vectors $(p_i^1(t), p_i^2(t))$, that is, the probability of player i selecting the strategy j (1 or 2) at period t.

We may depict the game (choice scenario) between two players by representing the reward matrix of the game structure as the matrix $ABCD$, the computation of the payoffs (π_1^j, π_2^j) at period t for a game as:

$$\pi_1^j(t), \pi_2^j(t) = \begin{pmatrix} p_2^1(t), \\ p_2^2(t) \end{pmatrix} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix} .(p_1^1(t), p_1^2(t)) \tag{2.7}$$

This approach returns the matrix cell containing the expected payoffs (π_1^j, π_2^j) for players ROW and COL, respectively, at time t, which we will denote throughout the analysis as *utility*, for the sake of simplicity. The fitness (f) for a game (g) in the population (G) was then defined as:

$$f_g(t) = (\pi_1^j(t) + \pi_2^j(t))^2. \tag{2.8}$$

The acquired utilities were aggregated to represent the overall utility derived from a game, then squared to give higher weights to the best-performing games in the game selection step relative to the current population's performance.

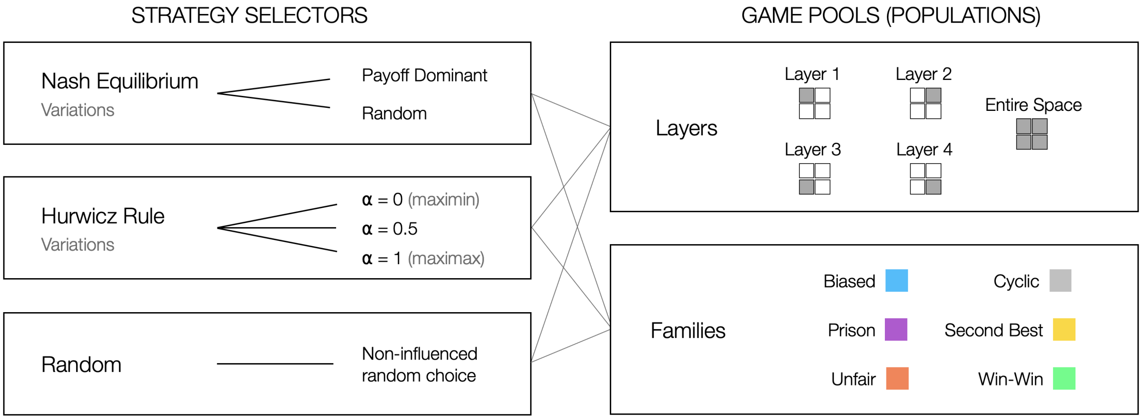


Figure 2.4: Overview – Strategy Selectors versus Populations

All variations of strategy selectors were applied to all game populations, as illustrated in Figure 2.4. Once the strategies were selected, and the utility scores were assigned to the entire set of games, the game population was ready to be processed by the Genetic Algorithm environment.

2.4.4 Genetic Algorithms Implementation: Genetic Operators

The application of the Genetic Algorithms method in this study was designed to model the collective learning process within a population of games. Each individual (game) represented a search point in the space of all potential solutions for the introduced problems and a potential temporal container of current knowledge regarding the laws of the environment. The process starts by initializing a population; then it evolves towards enhanced performance regions of the search space utilizing randomized crossover, mutation, and selection processes, or genetic operators (Back, 1996). The optimization process in this context aimed to adjust the payoff structures in the games, to disseminate the best outcomes for the defined strategy selection rules, as described by Haupt, Haupt, and Haupt (1998).

Selection and Replacement

Selection is a crucial factor for directing the search process toward better individuals. According to Reeves and Rowe (2002); Rowe (2015), the fitness function assigns positive scores to well-performing games in the search space. To avoid early convergence, we opted for a non-deterministic approach, meaning that all individuals are considered with a certain probability during the selection process. Based on the experiments with distinct methods, we adopted the fitness proportionate selection, where the probability of a game g being selected is based on its performance against the rest of the population: $\frac{f(g)}{F}$. Where f is the game's fitness scores and F is the total fitness of the current population (aggregated).

The applied replacement method was based on the Inverse Selection criterion (Rowe, 2015), which links directly to the fitness function and the previously described selection method. The standard replacement rate chosen for the algorithm was 1. Since the population size remained constant, another game required removal whenever a new game was added to the current population. This rule made the poorer performing solutions more likely to be replaced than the better ones since the fitness determined the probability of being replaced.

Reproduction: Crossover and Mutation

The chosen reproduction method was a single-point crossover (Holland, 1992; Lucasius & Kateman, 1993) that chooses a random index position within the individual's binary structure, and the parts of the two parents were exchanged at this point - generating a new individual, or offspring, as demonstrated in Figure 2.5. The idea was to recombine building blocks (schemes) on different strings. The crossover operator ensured that new individuals inherited the parents' characteristics, likely to be high performers among the population.

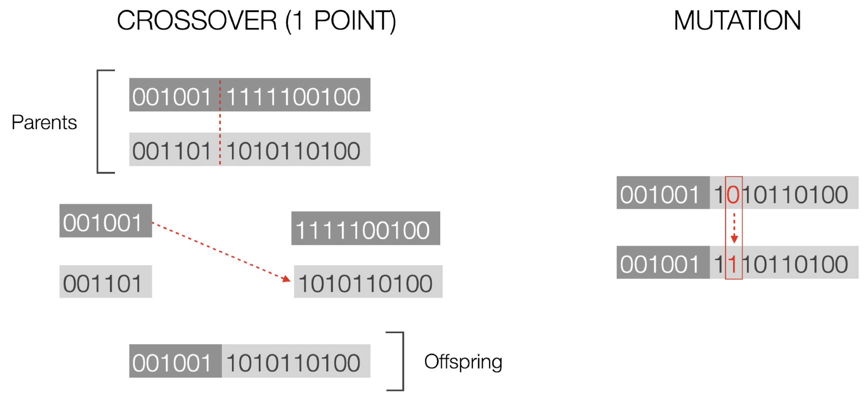


Figure 2.5: Crossover and Mutation Operators Example

As the next step, the mutation operator introduced a probability of changing one bit within the binary structure of an offspring game at a given generation. As illustrated in Figure 2.5 (Right-hand side), if mutation takes place, one random bit will be selected within the offspring game's (generated via crossover) binary string, with a uniform probability distribution over the possible 16 bits. The selected bit element is then flipped, from 1 to 0, or 0 to 1, generating a new game with new attributes, potentially not present in the parents' original gene pools. This step ensures the diversity of characteristics in the population and the possibility of generating more diverse games while avoiding fixation in determined sets of characteristics.

Evolutionary Loop

The outlined steps are iterated through multiple generations. The populations of games and their fitness scores were the main object of analysis for accessing the algorithm's properties and effects. The final data output consisted of a population of games with the same length as the initial set, outlining the effects of N generations of simulated natural selection, breeding, and mutation. Consequently, such generations directed the population to a convergence process towards an optimal point expected to become constant when first reached, meaning that the algorithm found an optimal solution for the problem. As a recap, the model applied the following steps:

1. initialization of a pre-filtered population of 2×2 strategic form games

2. Implementation of the strategy selection methods outlined previously

3. Calculation of the fitness for each individual (game) in the population, based on the strategies performance (fitness proportionate selection).

4. Execute the genetic operators: selection, crossover, and probably mutation within selected individuals in order to create a new game from the selected games.

5. Insert the offspring back into the population, replacing a game by inverse selection.

6. Return to step 2 and iterate this process until the termination criterion is satisfied

One important remark is that we employed a unified termination criterion for simplicity and experimental purposes at $10,000$ generations. Theoretically, the algorithm would have served its purpose when the entire game population reached the maximum fitness level (global optimum). We have recorded and documented the generation number in which all populations reach an optimal point as a measure of performance (speed of convergence) in table 2.3. The shared termination threshold allowed straightforward comparisons of the datasets and the observation of the stability of equilibrium states over more extended periods. Similarly, for the mutation rate, we have applied a unified parameter value at 1%. The mutation rate affects the evolutionary process and speed of convergence. Since we analyzed several different combinations of populations and strategy selection rules, we kept a constant parameter for simplicity and comparability. However, we encourage practitioners to find optimal mutation rates for individual problems in optimization-based tasks through parameter fitting or experimentation. We have performed a few tests with different mutation rates in terms of convergence. An example of such tests is found in Appendix 2.7.4.

Parameter	**Value**	**Short Description**
Distinct populations	11	Populations restricted according to layers and families logic
Strategy selectors	3	Nash equilibrium, Hurwicz rule and Random
Selection rule	FPS	Fitness Proportionate Selection (roulette wheel)
Crossover points	1	Random point where the binary string is divided and recombined
Mutation rate	1%	Probability of mutation taking place
Replacement rate	1	Number of games replaced by the offspring in each generation
Termination	10000	Maximum number of generations (iterations) allowed

Table 2.2: Summary: algorithmic parameters adopted in the model execution

Apart from the termination criterion situation described above, the simulation model comes with a range of other distinct parameters to be defined by the practitioner. Table 2.2 summarizes the methods and parameters applied in this experiment.

2.5 Simulation Results

Variations of both the Nash equilibrium and the Hurwicz Rule strategy selectors represented this study's concept of strategic behavior in decision-making. Alternatively, the absence of any strategic behavior was simulated by allowing the agents to select random strategies from the available options utilized as a control feature. The analysis of the experiment findings focused primarily on studying the effect of strategic behavior during the evolutionary learning process.

We quantified the influence of strategic behavior as the ability and speed of converging to optimal outcomes for an entire population, expressed by the individual utility levels. We applied the variations of the strategy selection methods into a range of different game pools, intending to assess this performance and also judging by the ability to transform the initial scenarios, which might be anything between pure conflict and harmony.

In addition to the utility convergence and strategy performance, the analysis of the simulation outcomes revealed two other interesting points, the coordination patterns in strategy selection by the agents and the transformation of the diverse game pools into sets of repeated games. Each of these points was explored more in-depth hereunder.

2.5.1 Utility Convergence

The simulation results indicated that the model mostly evolved in conformity with the reported literature and the expected development in accordance with the self-maximizing applied behavioral patterns. The evolutionary process drove the players into creating better outcomes throughout the learning generations, individually, by changing the structure of the games they were inserted into. The overall tendency was that once an optimal strategy was found, there were no incentives for the participants to deviate. Over time, successful tactics were generally replicated, resulting in an evolutionary equilibrium, as described in Reschke et al. (2001); Riechmann (1998, 2002). When the equilibrium was disturbed by the stochastic mechanisms in place, the players tended to evolve again toward an optimal strategy. Only the rounds with the Nash and Hurwicz strategy selectors get the aforementioned results. The Random approach produced fuzzier and unordered payoffs, which were usually non-optimal.

The analysis of the utility development, that is, the payoffs gained as results of strategies selected, was plotted in Figures 2.6, 2.7 and 2.8, for each of the populations and strategy selectors. Each plot's data is divided into two combined plots displaying the utility development for each generic player type, ROW, and COL. The Y axes show the average utility levels for the whole population in each generation (X axes), using different colors for each population. Each figure shows the results for each of the strategy selector types, being Figure 2.6 for the Nash equilibrium-based methods, Figure 2.7 for the variations of the Hurwicz rule, and Figure 2.8 for the completely random method. The populations are further divided into plots using the layers and family groups described earlier.

The evolutionary process gradually increased the individual utility to optimal values. The agents proceeded to select strategies that maximized the individual payoffs to maximize the overall outcomes of the games themselves, as defined by the fitness function. In this manner, the equilibrium became stable when all games in the given population achieved optimal payoffs. The selected strategies tended to persist and resist other invading strategies, as also observed by D. Friedman (1991). The charts displayed similar convergence curves in most cases.

Due to the stochastic mechanisms inherent to Genetic Algorithms, some games in the population may have lost an optimal pair of strategies during the evolutionary process. Nevertheless, the subsequent generations quickly adapted to the optimal strategies again. Similar effects of genetic experiments are observed in Riechmann (1998, 2001, 2002), where the learning process was given in two states: (1) the movement of populations towards a stable state, denominated behavioral stability and (2), once such state has been reached, the learning process presents a near-equilibrium dynamic of getting out of the evolutionarily stable state and returning there again. We observed similar trends in the equilibrium states observed in our results.

When looking at the Nash equilibrium runs (Figure 2.6), one can observe that in all game pools, the Nash strategies could drive the convergence to the optimal utility values. Interestingly, the population that took the longest to achieve the equilibrium was *Layer 2* (left side), which contained the most *Biased* games, mixed with *Unfair*, *Second Best* and *Prison* games. Similarly, when compared to the population containing all *Biased* games (right side), they also displayed a lower convergence time, even when compared with games having initially inferior strategies. Analysis inferences from the random Nash equilibrium strategy selection, apart from the inferior performance to the payoff maximizing version (as expected), include the fact that the more diverse pools (layers division) presented a better performance than the pools with all similar games together (families division). In this case, the pools with *Biased*, *Unfair*, and *Prison* games displayed a significantly higher convergence time.

Following data analyses of the populations using the Hurwicz rule, the equilibrium state was reached at similar speeds (see Figure 2.7) compared to the Nash equilibrium runs. According to the simulation results, the layers-based families presented similar effects as before, with *Layers 1* and *2* being slower than the others in some cases. There was a more distinct evolutionary pattern in the families-based populations until reaching stability. The *Second Best* and *Prison* families demonstrated a much noisier and longer process to reach equilibrium when considering the optimistic and, even more so, the pessimistic approaches.

The random pools displayed an interesting control feature, exhibiting a noisier process, with marked lower populations that could not converge to the evolutionary stability, even after ten thousand iterations, for both layers- and families-based populations (Figure 2.8). This lack of convergence properties is an essential contrast to the other runs, implying that the presence of strategic behavior highly influenced the evolutionary process itself. Even with a model tailored to favor individual payoffs, it was insufficient to drive the equilibrium state.

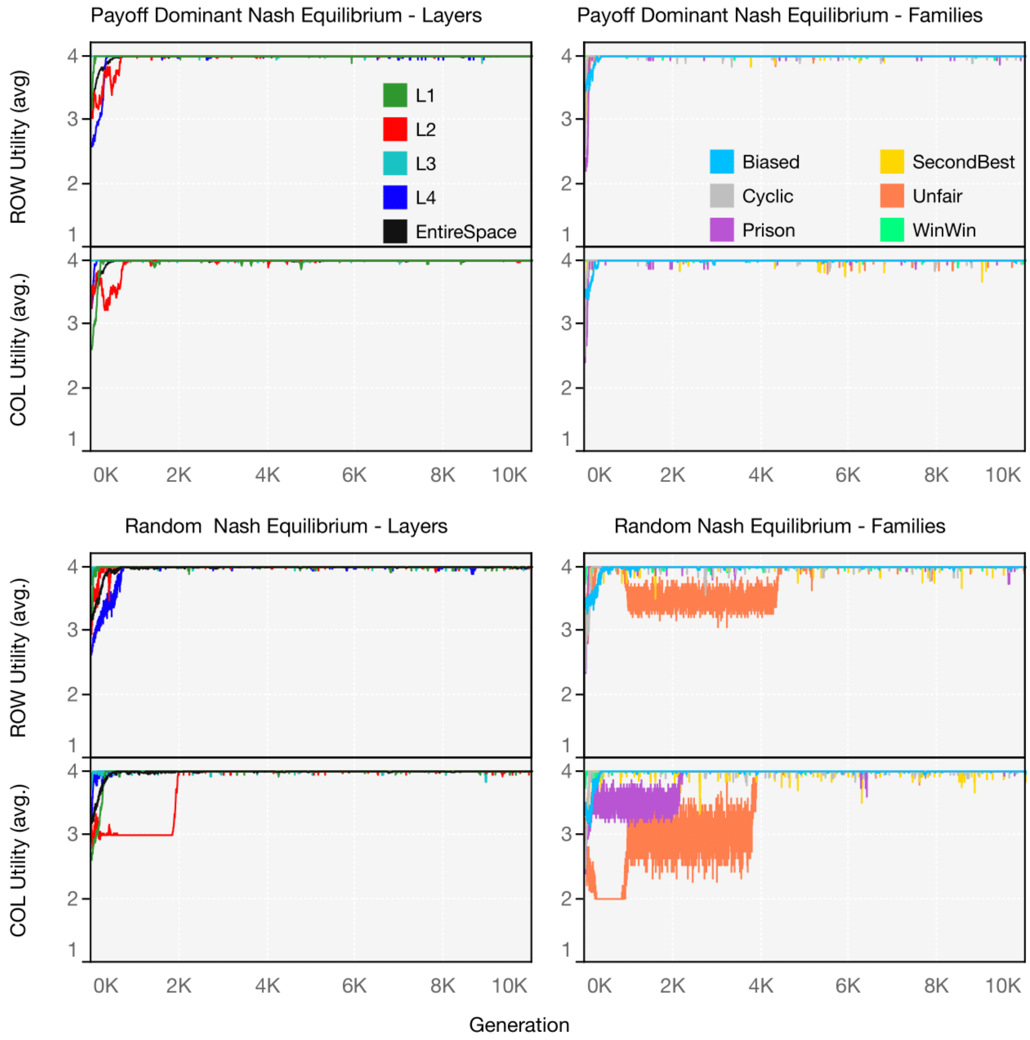


Figure 2.6: Utility Development - Nash equilibrium Strategies

2.5.2 Strategies Performance

This section analyses the strategy selectors' performance in terms of convergence speed to optimal values. Table 2.3 summarizes the speed of convergence results, that is, the number of generations it took for the entire population to reach maximum fitness scores. The payoff dominant Nash equilibrium strategy displays the best overall performance, followed by the Hurwicz rule with $\alpha = 0.5$, which alternates between the maximin and maximax strategies according to the game structures. In the third place, we have the Hurwicz rule with $\alpha = 1$, or maximax (optimistic), followed by the Nash equilibrium with random equilibrium selection, and the Hurwicz rule with $\alpha = 0$, or maximin (pessimistic). As denoted in the previous chapter, one can notice the lack of convergence for the random strategy selection method. In most cases, the random selection failed to converge within the allowed time, reinforcing the influence of strategic behavior simulation in the process.

Upon analyzing the game pools separately, we observed that pools with a higher number of conflict and non-optimal types of games, such as *Biased*, *Unfair*, *Second-Best* and *Prison*, affect the speed of convergence as well, requiring more iterations in order to reach opti-

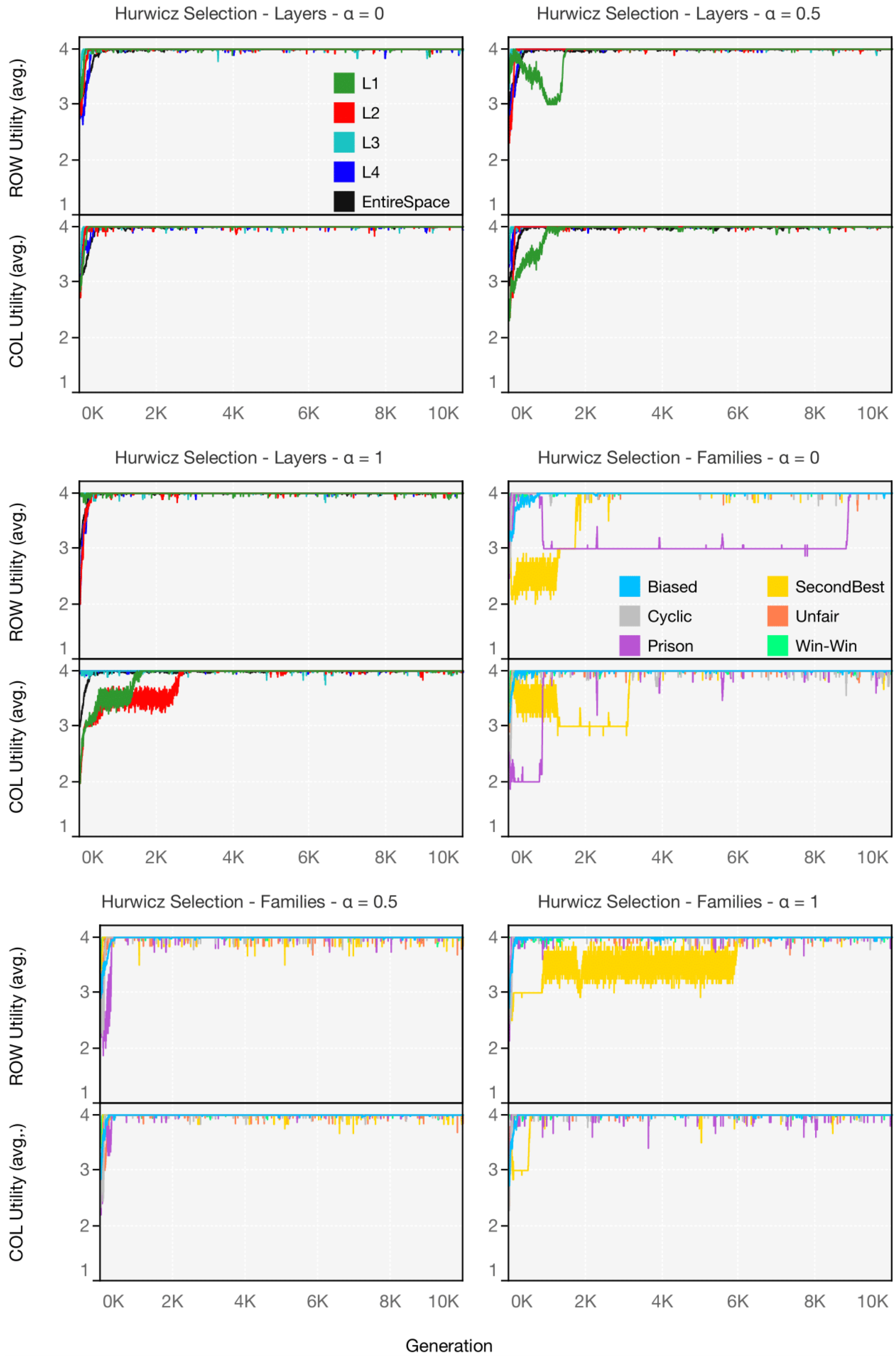


Figure 2.7: Utility Development - Hurwicz rule Strategies

mized fitness values. Although the games will be transformed in every case and eventually

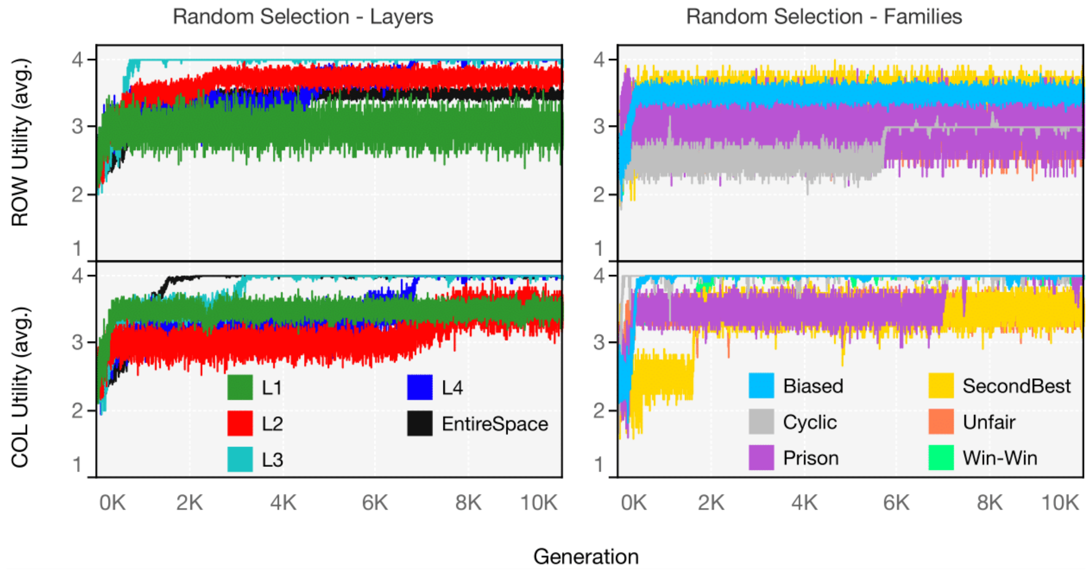


Figure 2.8: Utility Development - Random Strategies

converge to optimal scenarios when in the presence of strategic behavior, the available population characteristics directly influence how the games will transform in the next generations.

Population	**NE: pdom eq.**	**NE: rand eq.**	**HR** $\alpha = 0$	**HR** $\alpha = 0.5$	**HR** $\alpha = 1$	**Random**
Biased	363	401	737	392	217	failed
Cyclic	70	149	90	161	89	failed
Prison	95	2160	8881	298	54	failed
SecondBest	53	101	3148	55	4549	failed
Unfair	116	4382	73	235	58	failed
Win-Win	1	27	33	40	1	failed
L1	208	316	226	1494	1491	failed
L2	828	1972	209	178	2600	failed
L3	1	23	81	39	1	3201
L4	355	681	432	319	271	6842
Entire Space	701	653	742	709	612	failed
Average	254	988	1332	356	904	n.a.
Min	1	23	33	39	1	n.a.
Max	828	4382	8881	1494	4549	n.a.
Std. Deviation	283	1349	2658	424	1456	n.a.
Rank (avg.)	**1**	**4**	**5**	**2**	**3**	**6**

Table 2.3: Rank of strategies performance in each of the distinct populations

The charts in Figure 2.9 demonstrate the utility development for both players by adopting each strategy selector and variations used in the simulation model, aggregated by average across all populations. Here it is easy to understand this rank and compare the random method against the others. The performance of the strategies varied on a game basis; a complete overview of each combination of game pool and strategy selector is presented in Appendix 2.7.1.

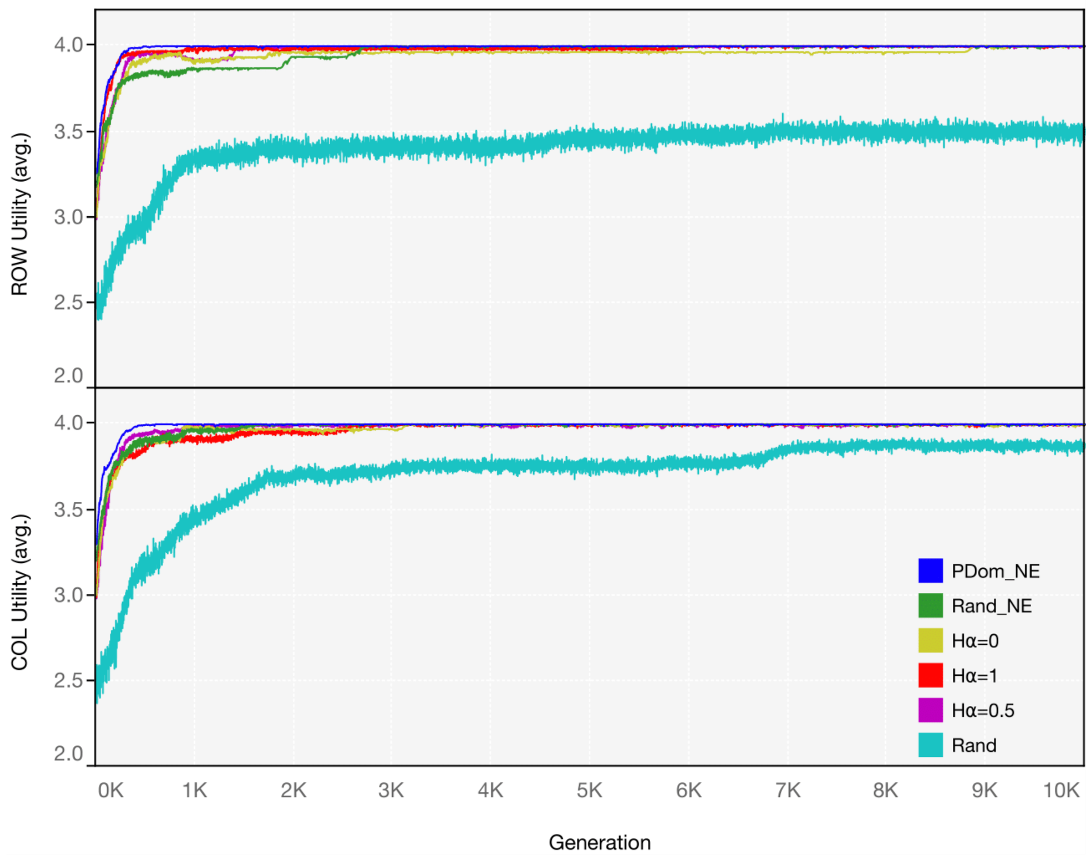


Figure 2.9: Strategies Performance - Consolidated

2.5.3 Evolutionary Stability: Coordination

Following our game structure, each agent could select between two strategies, generically denominated as *Strategy*1 and *Strategy*2. Additionally, for the Nash equilibrium-based pools, mixed strategies were allowed. The frequency of selected strategies based on a random sample of populations is presented in Figure 2.10. One can notice that the evolutionary stability, in terms of utility, is reflected directly in the stability of the strategies adopted by each player. This observed trend was in line with the examples detailed in Hines (1987); Weibull (1997). Furthermore, most displayed a symmetric plot, especially after reaching equilibrium.

The literature suggests that in equilibrium situations, the players coordinate their strategies. For each strategy one of the players adopts, there is a strategy that is always the best response, giving no incentives for the players to deviate from this equilibrium state by selecting another strategy. This allows the equilibrium state to persist throughout the generations (see Figure 2.10).

In addition, due to their inferior outcomes, mixed strategies have been eliminated early in the evolutionary process. Figure 2.10 displays a random sample of six strategy selectors and game pool combinations. The complete overview of the frequency of the selected strategies in all pools is presented in Appendix 2.7.2.

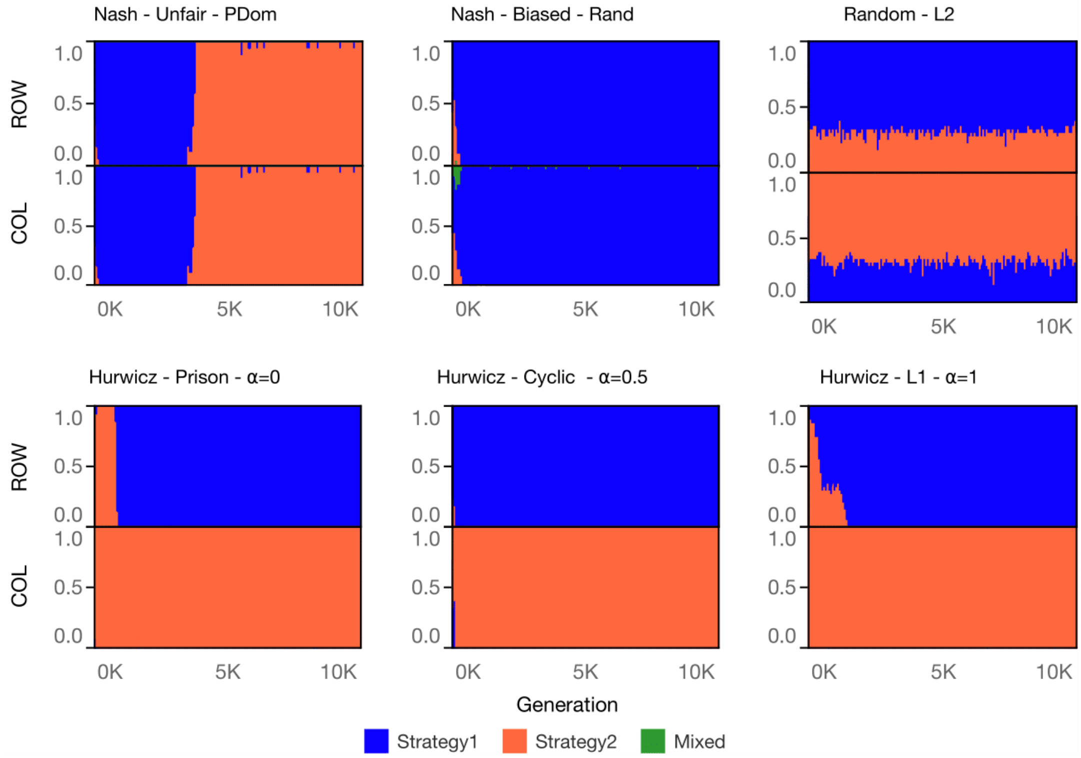


Figure 2.10: Evolutionary Stability - Sample Pools (percentage of strategies used in each round)

2.5.4 Transformation of Games

We started with a game pool containing the following equilibrium structure: one pure Nash equilibrium (75% of the games), one completely mixed Nash equilibrium (12.5% of the games), and two pure and mixed (one or infinitely many) Nash equilibrium (12.5% of the games). Table 2.4, depicts the resulting equilibria structure, depending on the strategy selection method, of the new games generated by the evolutionary process, considering the games in the last period of the simulation rounds.

Equilibrium Structure	Initial Pools	Hurwicz Pools	Nash Pools	Random Pools
1 pure	**75.0%**	74.8%	48.7%	0.0%
1 completely mixed	**12.5%**	0.0%	0.0%	0.0%
2 pure (and mixed)	**12.5%**	25.2%	51.3%	100.0%
Total	**100.0%**	100.0%	100.0%	100.0%

Table 2.4: Frequencies of the equilibrium structure for games in the final populations, compared to the initial pools

Games with completely mixed Nash equilibria have been eliminated during the evolutionary process, making all new games with at least one pure Nash equilibrium. The symmetry in the payoff structures was only sometimes present. A highlighted finding was the existence of an optimal equilibrium cell $(4, 4)$, where both players attained the maximum payoffs by playing that strategy, as in the Win-Win (harmonious) game family.

The analysis of the results uncovered another interesting fact. Most games transformed and replicated one-selves, reducing the number of unique games in the final populations, where successful games appeared repeatedly. There was no restriction to how the game may change, except for the rules defining the periodic table of 2 × 2 games (Robinson & Goforth, 2005). Even when the payoff symmetry was broken, both players had an optimal equilibrium state, which tended to survive across generations. The games were changed in such a way that they retained (in the majority of cases) the strategy choices, and payoff yield stabilized according to the characteristics of the initial populations and the decision rules applied. In this case, the games were specifically optimized to create favorable decision scenarios for both players.

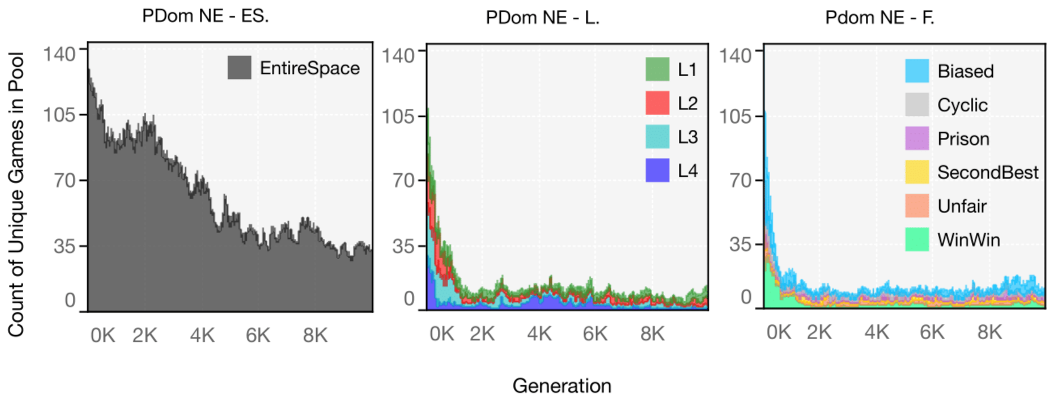


Figure 2.11: Count of Unique Games in Pool Across Generations - Sample Pools

This point is visualized in Figure 2.11, which contains a random sample of pools plotting the count of unique games in each of the denoted populations. This reduction pattern was equivalent across every population and strategy selector. The model yielded populations with few different games, repeatedly occurring within the same pool. The complete overview of the count of games in all pools is found in Appendix 2.7.3.

2.5.5 Comparison with Reinforcement Learning

Reinforcement Learning is a broadly applied concept in the economics literature, especially when it comes to game theory and situations of strategic interaction, as a model of the human learning process, as it is regarded as an alternative method to this paper for modeling learning behavior. A popular version of the model was introduced by Erev and Roth (1998), in which payers change their choice probabilities in reaction to payoffs from previous rounds. Duffy (2006) documented comparisons between Genetic Algorithms and Reinforcement Learning in the context of economics by mentioning two examples. In the first comparison, Haruvy and Ünver (2002) analyzes a scenario based on a procurement-type market experiment, where buyers and sellers are supposed to reach a stable outcome. Applying Reinforcement Learning and Genetic Algorithms yielded similar predictions consistent with empirical evidence. In the second example, Arifovic and Ledyard (2004) compare both methods' performances in the context of a public goods game, also employing a reinforcement-belief hybrid model known as Experience-Weighted Attraction (C. Camerer

& Hua Ho, 1999). By measuring performance in terms of time taken to converge to an equilibrium state (same rule used in our paper), the authors conclude that Reinforcement Learning ranks significantly worse than the other two methods. In contrast, the Genetic Algorithm displays the best performance.

As the model introduced in this paper is highly customized, a comparison might be beneficial for the reader to understand the discussion about performance and speed of convergence. For this reason, we applied a version of Erev and Roth (1998) 's model to the same pools of games and identically documented the results in terms of utility optimization and strategy selection. However, it is of uttermost importance to understand that in our original model, we allow the Genetic Algorithm to modify the games in a population. In contrast, the Reinforcement Learning model can only learn optimal strategies restricted by the static structure of the games, which will not change over time. In other words, as the programs start iterating through the data, they will eventually be processing different sets of games, even though the starting pools are the same.

Details on how we implemented the Reinforcement Learning algorithm can be found in Appendix 2.7.5. The model implemented for the comparison takes two parameters for a better fit to the data: ϕ, denoting the "forgetting rate," that is, how quickly the agents forget past payoffs, and λ, which defines the sensitivity to the weights assigned to strategies for the generation of choice probabilities. For simplicity, we applied static parameter values for all game pools, found through over 1000 simulations of 1000 rounds each. We calculated the average values of the pair of parameters that were most frequent in optimized outcomes in each population. Based on the results, the global parameter values were set at $\phi = 0.42$ and $\lambda = 0.9$, and initial attractions were set to 0; that is, no previous experience is assumed for the agents in any of the games.

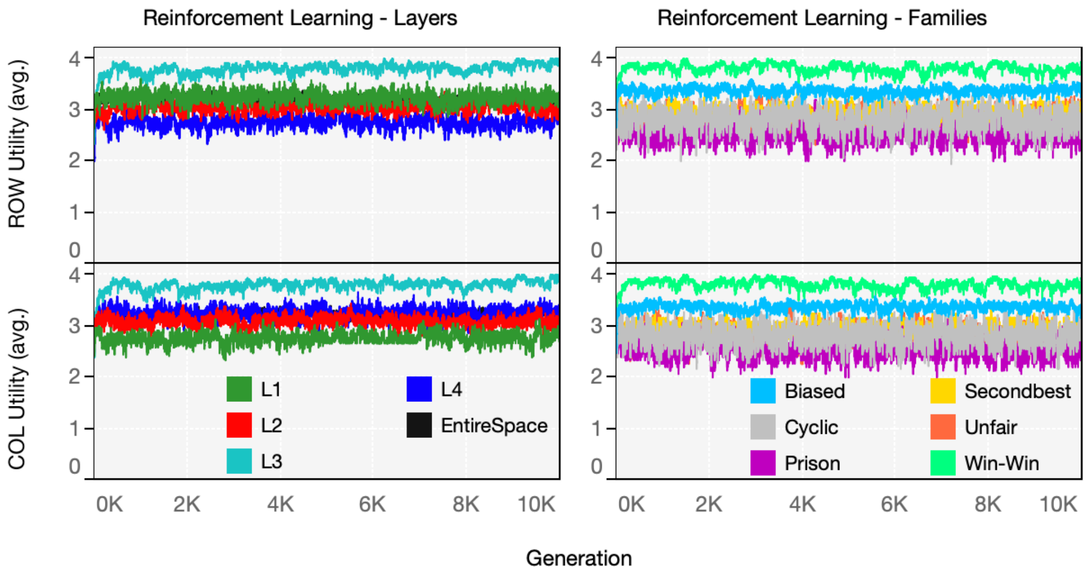


Figure 2.12: Reinforcement Learning Algorithm Application

Figure 2.12 displays the results for the Reinforcement Learning application. As suggested by earlier comparisons, Reinforcement Learning (and many other similar models), by nature, find suitable strategies that maximize the rewards in a given scenario. The fitness values converged to local optimum values. However, the main difference in our setting is that the Reinforcement learning model is constrained to the structure of the games, as observed in the summary of the results. The utility will never reach a global optimum if optimal-yielding strategies are not available for selection, as the games remain static over time.

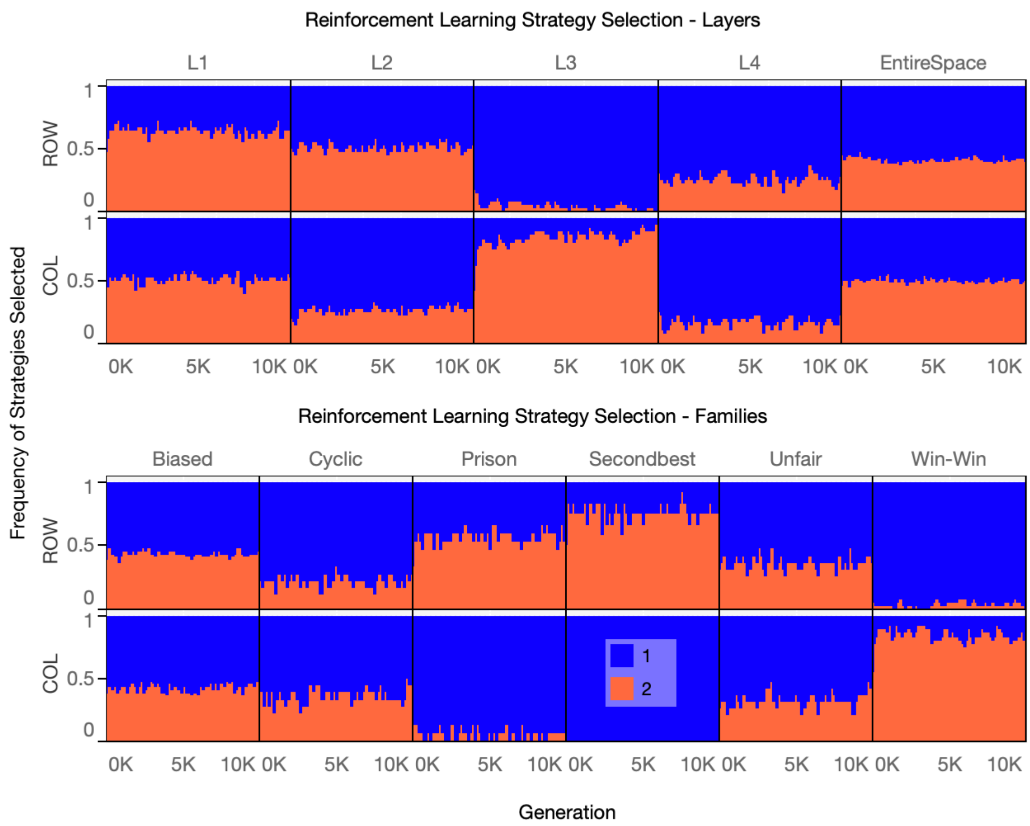


Figure 2.13: Reinforcement Learning Strategies Selection Overview

When analyzing the strategy selection behavior of the agents, we noticed weaker coordination of strategies using reinforcement learning, as demonstrated in Figure 2.13. Agents learned optimal strategies within each of the games. However, no consistent optimal-yielding strategies can be selected across an entire population, outlining the Genetic Algorithm's capacity to optimize multiple scenarios by transforming individual situations. Therefore, this higher coordination effect cannot be achieved in optimizing multiple static games, and it shows similar figures to the random strategy selection method introduced before within the Genetic Algorithm implementation.

Fundamentally, both methods are based on similar premises but have different operating processes. Reinforcement Learning methods, such as the one used for this comparison, proved to be very efficient in finding optimal payoff-maximizing strategies, even using global parameters. However, the Genetic Algorithm performs better in our environment as strategic behavior rules allow the algorithm to change the games based on well-performing strategies. This ability enables the transformation of the games based on the agents' behavior, which optimizes the decision environment as a whole, introducing the possibility of new, potentially better strategies.

2.6 Conclusion and Discussion

This paper introduced a novel game-theoretical approach for analyzing dynamic scenarios that uses a heuristic approach to transform games played sequentially based on the decisions performed by the agents, taking inspiration from observed scenario transformations over time. Based on how economic agents learn and transform their decision environments in the real world, our motivation was to document and describe the process of implementing a Genetic Algorithm as a mechanism for evolutionary economic learning. For this purpose, we compared the convergence rate to optimal scenarios of different strategy selection rules in different populations of diverse 2×2 game types. The presented analysis demonstrated how strategic behavior can influence this dynamic learning process and how different strategic behavior mechanisms perform.

The core observations extracted from the presented results are summarized in the following points:

1. The simulated strategic behavior for decision-making directly impacted the evolutionary process, being responsible for transforming games into optimal scenarios (from conflict to harmony).
2. Similarly, the absence of strategic behavior negatively affects the Evolutionary process, preventing the evolution toward optimized scenarios.
3. The rank of the strategy selection methods from best to worse performance is given as follows: (1) payoff dominant Nash equilibrium, (2) Hurwicz rule with $\alpha = 0.5$, (3) Hurwicz rule with $\alpha = 1$, (4) Nash equilibrium with random equilibrium selection, (5) Hurwicz rule with $\alpha = 0$ and (6) random strategy selection.
4. When agents maximize their utility (i.e., behave strategically), the games evolve to have higher payoff structures, optimizing their decision environments and transforming conflict into harmonious games.
5. The evolution of dynamic strategic situations tends to create games where coordination is possible.
6. Decision rules and the environment characteristics influence the optimization process's agility in reaching an evolutionarily stable equilibrium.
7. Comparisons with reinforcement learning outline the power of transforming the games for the evolution of conflict to optimal scenarios. The reinforcement learning model can find optimal strategies within the game constraints but cannot optimize further without structural changes in the population of games.

The diverse set of games used for this analysis provided a rich representation of multiple real-life scenarios, including conflict, biased, dilemma, and harmonious situations. In addition, combining multiple game types with multiple behavior types yielded a diverse data set for exploring additional individual factors, such as which strategies performed better and which games were the most challenging to achieve a stable equilibrium.

The introduced decision rules presented a satisfactory performance in supporting the transformation of the games in maximizing individual payoffs. An exception was the random selection method, which failed to reach the evolutionary equilibrium multiple times in the allowed time. In many points, the performance of the Nash equilibrium and Hurwicz rule variations ranked similarly. However, in essence, the Nash equilibrium was still found to be, in this study, the most rapid and robust method for optimization modeling, in conformity with past findings by Sefrioui and Perlaux (2000).

The self-maximizing behavior and the transformation of the games allowed the individual strategies to evolve so that the agents tended to coordinate their strategies. Hence, for each player's strategies, an optimal best response is consistent over time. When the genetic process eliminates stable strategies, the evolutionary learning process eventually creates new optimal strategies that enable a new equilibrium state, reinforcing the findings in Chmura, Goerg, and Selten (2012); Kalai and Lehrer (1993), the rational learning in replicated games eventually leads to stationary points of the Nash equilibrium. Such a result is also in line with the concept of genetic stability documented in Riechmann (1998, 2001, 2002). In essence, strategic behavior allowed the agents to transform each game scenario into mutually optimal situations, eliminating conflicts and inequality.

Per Savin et al. (2018), incorporating the dynamics of strategic scenarios in equilibrium analysis allows for more realistic interpretations of possible environmental transformations and policy outcomes. Compared to this paper, our results relied on transforming strategic scenarios due to strategic behavior. We also conclude that heuristic techniques and their extensions provide flexible tools with broader search spaces, potentially resulting in superior and even innovative outcomes than traditional procedures. The comparison with Reinforcement Learning in our research also outlines the limitations of commonly employed methods, proposing new solutions to more flexible problems and providing insights into reality-inspired dynamics.

As discussed by Brown and Rosenthal (1990); Chmura et al. (2012); Erev and Roth (1998), the actual human behavior is at times not well predicted by the standard theory. However, in reality, the evolutionary social models are more complex and require significant efforts to create representative pictures that capture relevant characteristics of social systems (Reschke et al., 2001). The model created in this study aimed at an exploratory analysis that contributes to the discussion on how to envision and build representative models of strategic behavior and economic learning in different situations, exploring the effects of decisions in repeated strategic interaction.

We developed our simulation model to be flexible and integrate other ideas that enable the examination of dynamic games, strategic behavior, and economic learning in the future. Further research shall be performed by enriching the current model with different types of games (also diverse populations) and encoding different strategy selection models, which should capture more diverse behavioral profiles of decision-makers in varying contexts. In another exploration direction, practitioners can add another layer of processing payoffs and utilities by defining a spectrum of profiles based on social preferences, such as altruism, envy, fairness, and justice, defined by the agents' utility functions. Furthermore, different Evolutionary Programming methods can be compared with the results achieved with the Genetic Algorithms application.

2.7 Appendix

2.7.1 Complete diagram of strategies performance

Figure 2.14 contains the overview of the performance achieved by the different decision rules in every population, individually.

2.7.2 Complete diagram of evolutionarily stable strategies

Complete overview of the strategies selection frequency for the Nash equilibrium Pools in Figure 2.15

Complete overview of the strategies selection frequency for the Hurwicz Pools in Figure 2.16

Complete overview of the strategies selection frequency for the Random Pools in Figure 2.17

2.7.3 Complete diagram of games transformation

Figure 2.18 displays the overview of the count of unique games across generations in all pools, for all strategy selectors.

2.7.4 Mutation Rate Tests

Figure 2.19 summarizes the tests performed with mutation rates using the Payoff Dominant Nash Equilibrium strategy selector, which is the highest performing and most robust of our rules, applied in the whole pool of 144 games.

2.7.5 Reinforcement Learning Algorithm

The reinforcement learning algorithm applied for a comparison with the Genetic Algorithms was based on Erev and Roth (1998) 's model and in the formulation presented in C. Camerer and Hua Ho (1999); Moffatt (2020), which introduced the concept of attractions A_i^j, as weights attached to strategies. Initial attractions $A_i^j(0)$ can be given values to assume a degree of previous experience from the agents with the game (having to be estimated by the practitioner), or they can be neutral by setting default values to 0. The generalized model has two main components. First, the attractions' update function at time t $A_i^j(t)$, given as:

$$A_j^i(t) = \phi A_i^j(t-1) + I(s_i^j, s_i(t))\pi_i(s_i^j, s_{-i}(t)), \tag{2.9}$$

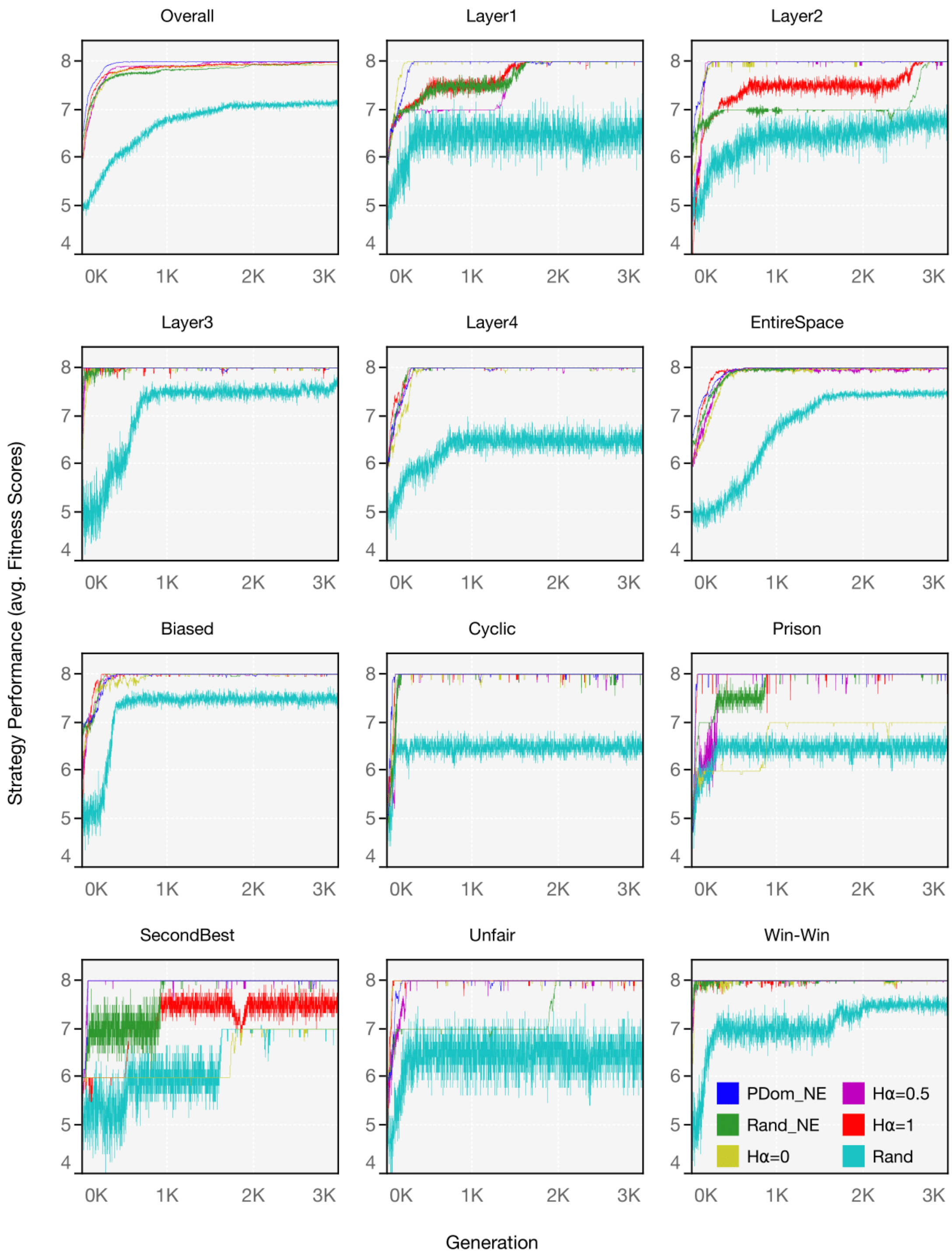


Figure 2.14: Strategies Performance - All Pools

where $\pi_i(s_i(t), s_{-i}(t))$ is the player i's payoff in round t (scalar-valued function) and I is the indication function, taking the value of 1 if the statement is true and 0 otherwise. This means that in reinforcement learning, a player's attraction to a strategy can only increase if that strategy is chosen. ϕ indicates the speed at which past payoffs are forgotten. $\phi = 0$ would indicate that only the most recent payoff is remembered. $\phi = 1$ would indicate that all past payoffs have equal weights in the current decision.

The second component is the transformation of attraction into choice probabilities via

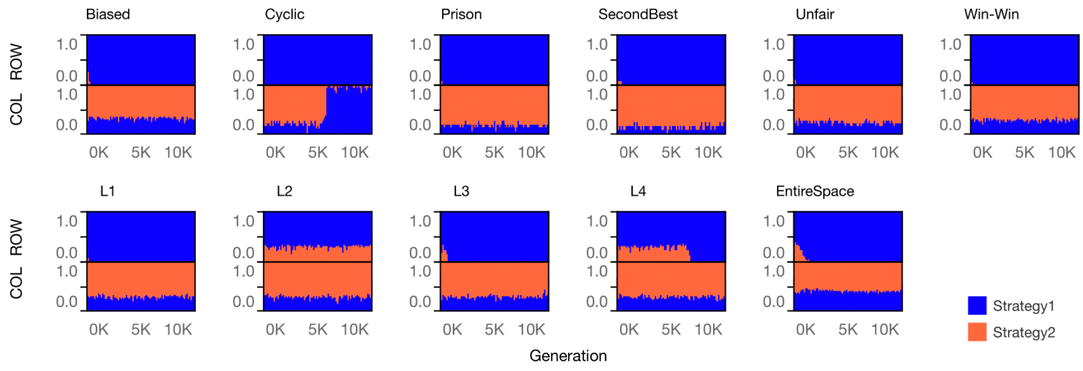


Figure 2.15: Evolutionary Stability - Nash Pools

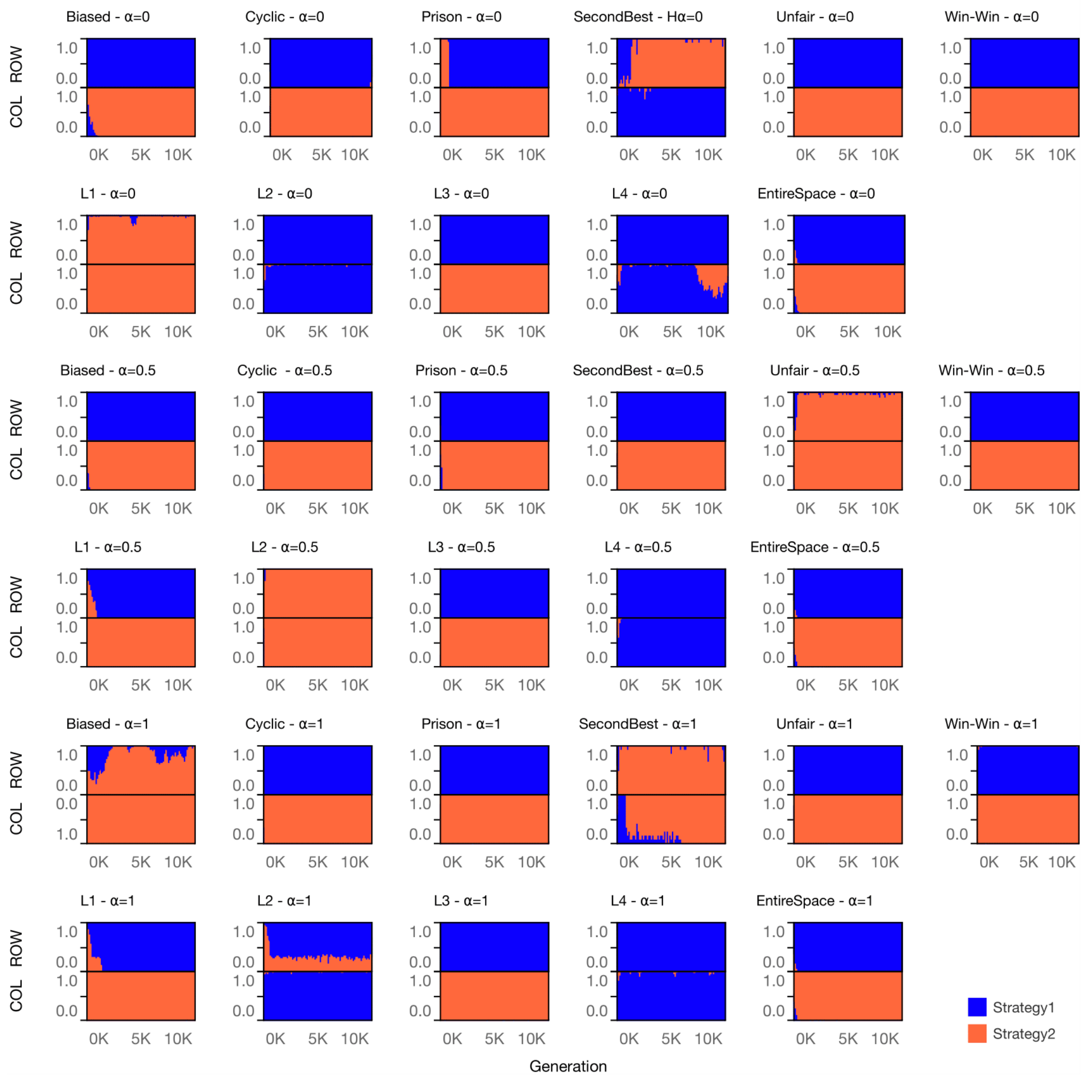


Figure 2.16: Evolutionary Stability - Hurwicz Pools

logistic transformation, given as:

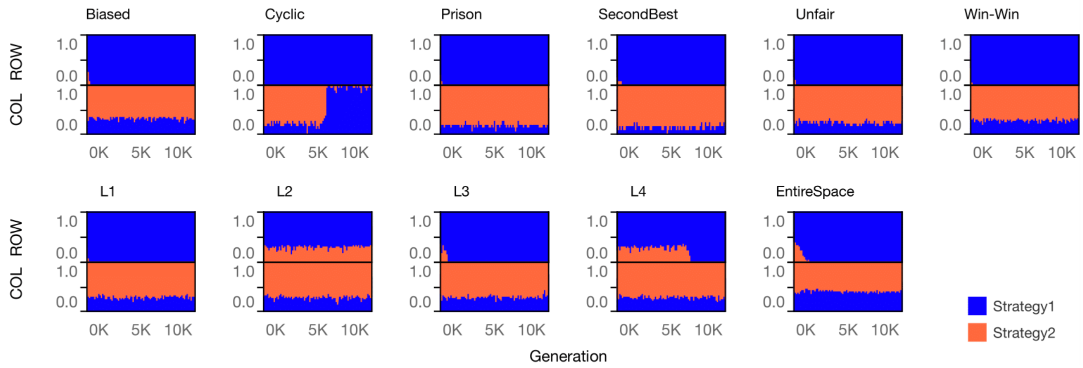


Figure 2.17: Evolutionary Stability - Random Pools

$$P_i^j(t+1) = \frac{e^{\lambda A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda A_j^k(t)}} \tag{2.10}$$

Where $P_i^j(t+1)$ is the probability of player i playing strategy j in round t and m_i is the number of strategies player i has. The parameter λ defines the sensitivity to attractions; attractions are irrelevant if $\lambda = 0$, and attractions are important if λ is large. The choice probabilities are then used to select strategies in each round.

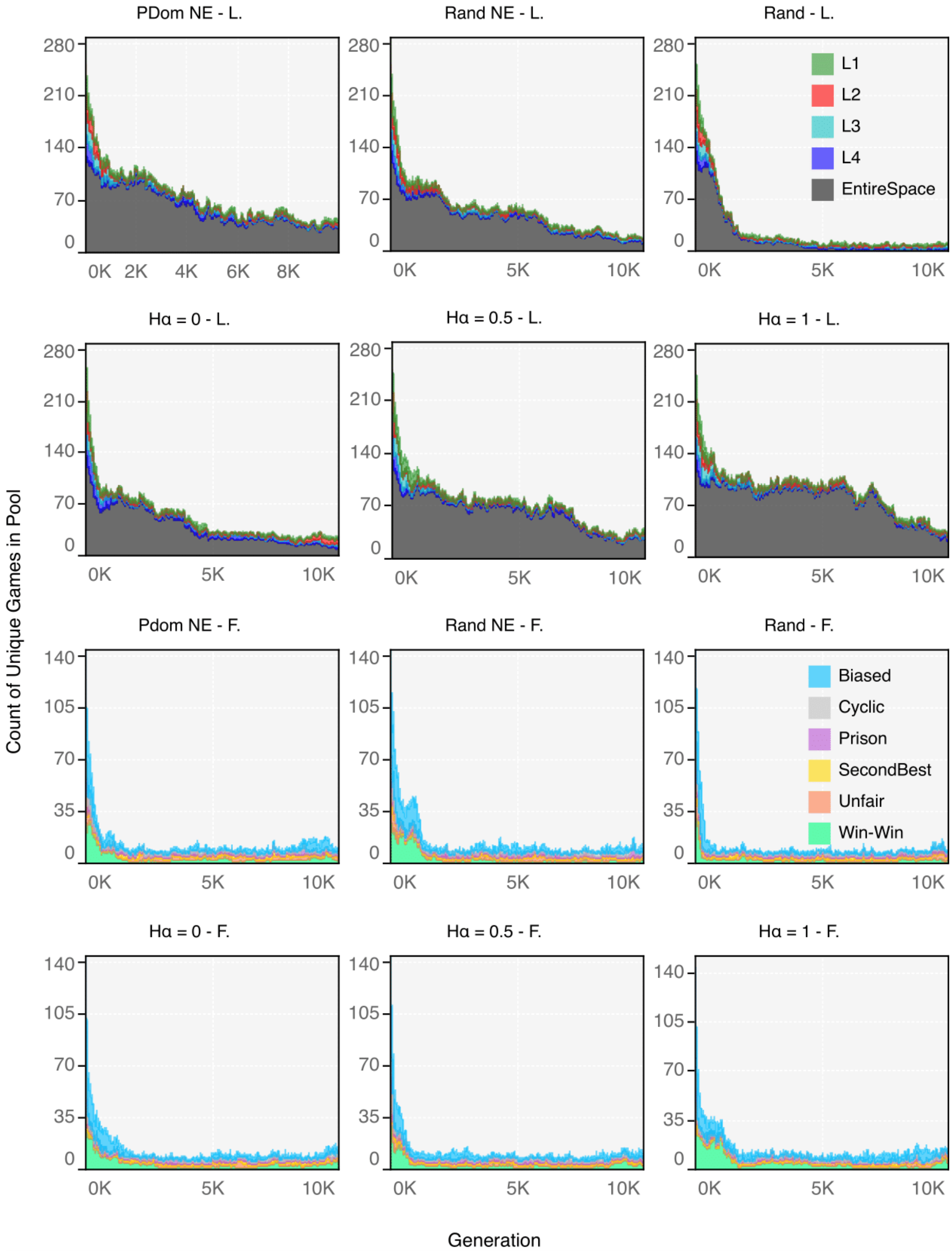


Figure 2.18: Count of Unique Games in Pool Across Generations - All Pools

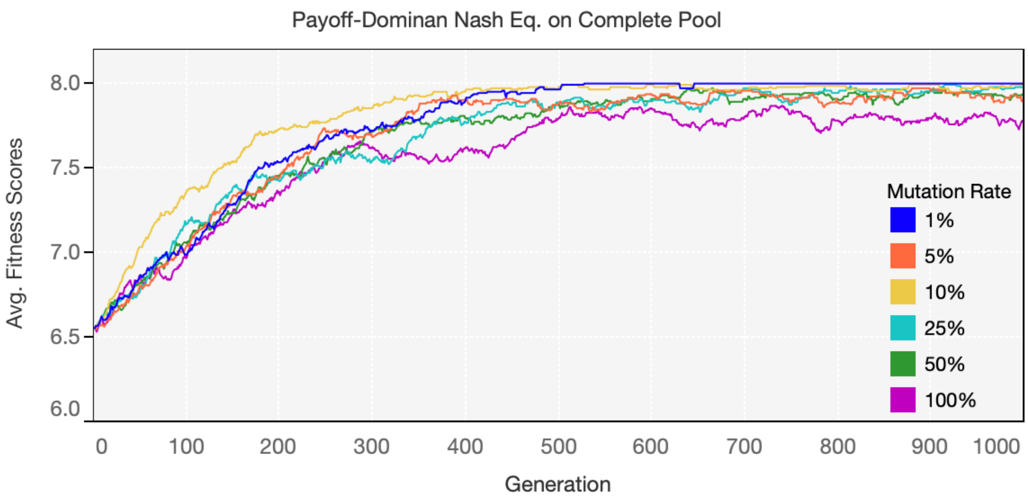


Figure 2.19: Mutation rate tests - Payoff-Dominant Nash Eq. applied on the complete pool of games

3. Stationary Equilibria in Behavioral Game Theory: An Experimental Analysis of Inspection Games

Authors

Vinícius Ferraz, Thomas Pitz, Jörn Sickmann, Wolf Gardian & Deniz Kayar

Abstract

This paper investigates the predictive capabilities of different stationary equilibrium concepts within the framework of an inspector game. The experiment employed a 2×2 asymmetric payoff design spanning 70 periods, executed under information-loaded and neutral frames. Drawing on data from 100 participants, we analyze the five established stationary equilibria concepts and five modified versions incorporating loss aversion and fairness parameters. Our analysis emphasizes predictive performance and model characteristics on aggregated, time series, and game-play data aggregations. The results show the limited predictive power of the Nash Equilibrium, while the Action Sampling Equilibrium and Impulse Balance Equilibrium emerge as the best predictors among the original concepts. The modified models exhibit high predictive potential but also increased calculation complexity and parameter estimation. The modified Impulse Balance Equilibrium with a dynamic loss aversion parameter stands out for its predictive power and robust representation of the loss-aversion behavioral trait.

Keywords

Behavioral Game theory, Equilibrium Theory, Stationary Equilibrium Concepts, Inspection Games, Competition

3.1 Introduction

Understanding the motivations behind equilibrium in strategic interactions gives us several insights into predicting or describing situations involving human decisions. While traditional game theory relies on strict rationality, behavioral game theory expands this by accounting for psychological effects, context, and other human factors (C. F. Camerer & Loewenstein, 2003). This paper introduces an experimental study that examines equilibria in a game that incorporates psychological elements like morality, retaliation, punishment, and cheating, modeled with an inspection game.

Inspection games depict scenarios where an inspector monitors another entity, termed the inspectee, ensuring adherence to specific rules (Avenhaus, Canty, Kilgour, Von Stengel, & Zamir, 1996). These models can shape the creation of optimal incentives, such as penalties or bonuses, to foster desired economic behaviors. Notably, games with entirely mixed strategies (as inspection games), where players employ every strategy at certain probabilities (Kaplansky, 1945, 1995), pose challenges to traditional rational models, such as the Nash Equilibrium (Erev & Roth, 1998)

Our study investigated a student versus inspector interaction. Drawing inspiration from Selten and Chmura (2008), we designed an experimental framework to evaluate five stationary equilibrium concepts: Nash Equilibrium, Quantal Response Equilibrium, Impulse-Balance Equilibrium, Action Sampling Equilibrium, and Payoff Sampling Equilibrium. Our analysis of the equilibrium behavior includes an in-depth investigation of the models' characteristics and performance in different levels of aggregation. In addition, we have also investigated whether framing in such games influences the participants' strategy selection by applying two versions of the same game: one neutral frame, where no information about the game is given, and an information-loaded frame, containing a contextual description.

Our analysis extended to examining model parameters by incorporating behavioral traits into predictions. We introduced five variations of the original stationary equilibria concepts, integrating loss aversion and inequity aversion elements. Our primary goal was to systematically evaluate how these modified concepts aligned with observed game equilibria and to understand their inherent features more comprehensively. In line with Mauersberger, Nagel, and Bühren (2020) suggestion that adding behavioral components to economic models can enhance our understanding of real-world phenomena—taking the 'level k' reasoning as an example—our approach seeks to enhance both predictive accuracy and explanatory power.

Our results demonstrated that we achieved enhanced predictive performance in inspection games by integrating behavioral traits such as loss aversion and inequity aversion into stationary equilibrium models. While traditional models like the Nash Equilibrium were challenged, modified versions provided more accurate predictions, especially when accounting for the temporal dynamics of player strategies. Interestingly, the influence of game framing was marginal, with more comprehensive information slightly inclining participants towards riskier strategies. Furthermore, our models' parameter investigation revealed that quantifying behaviors like loss aversion closely aligns with real-world player behavior.

This paper is structured as follows: we first outline the conceptual framework describing the key areas combined in this research. We then discuss the experimental design and framing techniques. The results section elaborates on the equilibrium analyses based on aggregated data, time series, and individual game levels. Finally, we evaluate the predictive capabilities of our model modifications by applying them to the other 12 cyclic games used in the experiments documented by Selten and Chmura (2008).

3.2 Theoretical and Conceptual Background

This chapter explores the foundational game theoretical concepts central to our research. We examine concepts such as inspection games, cyclic games, and various stationary equilibrium concepts in 2×2 games. As we navigate these ideas, we highlight how behavioral traits influence strategic decision-making, especially those related to fairness, inequity aversion, and loss aversion. The subsequent sections unpack these core concepts in detail.

3.2.1 Inspection Games

Inspection games model strategic interactions between two actor profiles: an inspector and the inspected. Within these games, the inspector must decide whether to oversee the inspected party, who in turn chooses between compliance and non-compliance with established rules. The associated payoffs and consequences drive their decisions. As a part of non-cooperative game theory, inspection games are recognized as discoordination games, wherein the optimal outcomes often emerge when both parties make contrasting choices. The foundational works by Dresher (1962) and Maschler (1966) provide a deep dive into its applications in warfare and economics. Subsequent research, such as by Kolokoltsov, Passi, and Yang (2013) and Nosenzo, Offerman, Sefton, and van der Veen (2016), has expanded its boundaries, highlighting evolutionary perspectives and mixed strategy equilibriums, respectively. For a comprehensive overview of the game's applications and mechanics, refer to Borch (1982), Avenhaus et al. (1996), and Avenhaus, Von Stengel, and Zamir (2002). Practical applications range from mitigating shirking, as evidenced by Nosenzo, Offerman, Sefton, and Van der Veen (2010); Nosenzo, Offerman, Sefton, and van der Veen (2014), to ticket controls and fraud detection in tax audits (Avenhaus, 2004).

3.2.2 Cyclic Games and Stationary Equilibria

The inspection games belong to the cyclic family, which describes the recurring nature of particular economic and social situations. In contrast to extensive form games, cyclic games might not have an end. The same scenarios can repeat infinitely often, with players who might enter and exit such situations at different times (Selten & Wooders, 2001). Gurvich, Karzanov, and Khachivan (1988) provided an early formalization of the concept and a procedure to find optimal strategies in a 2×2 setting based on the minimax concept.

In game theory, a stationary equilibrium denotes a scenario where players consistently adopt the same strategies throughout repeated game plays. This concept, highlighted by Selten and Chmura (2008), emerges from specific learning or evolutionary dynamics, offering a robust framework for analyzing repetitive game interactions. In this context, models of stationary equilibria aim to predict or describe the strategies that players will consistently adopt throughout repeated interaction periods. Selten and Wooders (2001) describes stationary equilibria as a form of local optimization: the strategies chosen at any decision point remain consistent, regardless of the sequence of prior decisions. This equilibrium is maintained as long as altering these strategies does not offer players a better payoff, given their opponents' decisions.

Selten and Chmura (2008) compared experimentally different stationary concepts for completely mixed 2×2 games, namely the Nash equilibrium, the quantal response equilibrium, the action-sampling equilibrium, the payoff-sampling equilibrium, and the impulse balance equilibrium. Concerning the predictive performance, the impulse balance equilibrium is the best, and the Nash equilibrium is the worst. Brunner, Camerer, and Goeree (2011) complement these studies; they also found that the Nash equilibrium concept fits the worst, while the action and payoff sampling concepts fit slightly better than the other non-Nash concepts.

Chmura, Goerg, and Selten (2014) reinforced that the generalized impulse balance equilibrium in 3×3 contexts offers predictions closer to the empirical data than the Nash equilibrium. Kirman, Laisney, and Pezanis-Christou (2018) compared impulse balance and quantal response concepts in binary choice participation games, assessing cluster heterogeneity consistency with symmetric models. Results showed better performance of impulse balance equilibrium over quantal response equilibrium in accommodating model-consistent cluster heterogeneity.

The literature suggests that impulse balance equilibrium may replace Nash equilibrium for cyclic game predictions, as it uses no parameters and offers better accuracy, assumptions discussed and tested in this study.

3.2.3 Stationary Equilibrium Concepts

This paper explores five such stationary equilibria that have been prominently discussed in academic literature. Additionally, we enhance these models by incorporating behavioral parameters, emphasizing loss and inequity aversion. This chapter is dedicated to providing clear, conceptual definitions for each of these equilibria methods. Please refer to appendix 3.6.2 for readers interested in detailed formal representations. For technical remarks regarding the implementation of the calculations, see appendix 3.6.3.

The Nash Equilibrium (NE), a foundational concept in game theory, defines a state where every player's strategy is the best response to the strategies chosen by other players. In this equilibrium, each player's expectations and strategies align perfectly. No player is incentivized to change their strategy unilaterally, given that others stick to theirs. It captures a state of mutual best response and strategic stability (Nash Jr, 1950). Although the NE is simple and does not require parameter estimations, Erev and Roth (1998); Selten and Chmura (2008) highlight that this method exhibits counter-intuitive features, rendering it a sub-optimal predictor for completely mixed games. We will explore this assumption further in later sections.

Taking the concept of bounded rationality into consideration, McKelvey and Palfrey (1995) introduced the Quantal Response Equilibrium (QRE) Concept. In QRE, players' decisions are influenced by the expected behaviors of others, but with a twist. Players do not always choose the absolute best response. Instead, their choices are probabilistic, with strategies that have higher expected payoffs being chosen more frequently but not always. This concept introduces a degree of "noise" or randomness into decision-making, reflecting the imperfections and uncertainties inherent in real-world choices. In contrast to the NE, the QRE introduces a parameter λ, which represents the precision of players' beliefs about others' strategies, with higher values indicating more accurate beliefs and more rational behavior. In other words, high values of λ converge QRE to NE.

In the Action-Sampling Equilibrium (ASE), based on Osborne and Rubinstein (2003), players do not respond to the overall strategy distribution of their opponents. Instead, they take a limited sample of the other players' recent actions and then decide on the best response to this sample. This approach introduces variability and acknowledges that players base their decisions on recent experiences or observations rather than broader strategic considerations. The sample size N for the opponents' actions is a parameter to be adjusted.

Similarly to the ASE, in the Payoff-Sampling Equilibrium (PSE), players are retrospective in their decision-making. They sample and assess the historical payoffs associated with each strategy. Players choose the strategy that has historically yielded better results by comparing these sampled outcomes. It is a method that emphasizes learning from past experiences and outcomes, as envisioned by Osborne and Rubinstein (1998). This concept also incorporates a parameter, N, representing the number of samples taken from one's payoffs.

Lastly, we transition to the Impulse Balance Equilibrium (IBE). Introduced in Ockenfels and Selten (2005), based on the learning direction theory (Selten & Buchta, 1994). In contrast to the previous concepts, the IBE diverges from the traditional best-response thinking. Instead, it considers the idea of "impulses" that push a player towards or away from a particular strategy. Central to the IBE is the concept of a "security level" or reference point, termed the pure-strategy *maximin*. Payoffs from each strategic choice are assessed against this reference, with gains halved in line with the principles of prospect theory (Kahneman & Tversky, 1979) idea of loss aversion, emphasizing that losses count twice as much as gains. This process results in a modified game matrix from which these impulses are derived. The equilibrium is achieved when the push and pull of these impulses balance out, leading to a state where players' actions are influenced by both their rational calculations and their emotional or psychological impulses.

The original IBE model lacks explicit parameters but incorporates an implicit constant for loss aversion. Brunner, Camerer, and Goeree (2010) highlight that the lack of parameters may simplify the model but limit its empirical application. Ockenfels and Selten (2005) and C. Camerer (2012) have also experimented with different loss aversion values. In our study, we introduce a dynamic variant, IBEγ, which treats this constant as a parameter γ and compares it with the original IBE.

3.2.4 Inequity Aversion Matrix Transformations

Broadening our investigation into how behavioral traits impact strategic equilibria, we have incorporated the notion of fairness through inequity aversion. This concept emphasizes an individual's distaste for inequitable outcomes. We have transformed the game matrix before computing the equilibrium using the utility function introduced in foundational work by Fehr and Schmidt (1999). Within this framework, we introduce two parameters: α, which represents a player's discomfort with earning less, often referred to as "envy," and β, which indicates a player's unease with earning more, known as "guilt." This adjustment renders the perceived payoffs dependent on fairness considerations (Tavoni, 2009). See appendix 3.6.2 for the formal description. We use $IA_{(\alpha,\beta)}$ as an abbreviation for the inequity aversion matrix transformations, generating the following modification for the NE and IBE:

- $NE \circ IA_{(\alpha,\beta)}$: Nash equilibrium with inequity aversion matrix transformation.
- $IBE \circ IA_{(\alpha,\beta)}$: Impulse balance equilibrium with inequity aversion matrix transformation, without the impulse transformation. This model replaces the aspiration level framework from the original model with the inequity aversion, as performed in Tavoni (2009).
- $IBE_2 \circ IA_{(\alpha,\beta)}$: Impulse balance equilibrium with both inequity aversion and impulse transformations, holding $\gamma = 2$, as in the standard model.
- $IBE\gamma \circ IA_{(\alpha,\beta)}$: Impulse balance equilibrium with both inequity aversion and impulse matrix transformations, with dynamic γ.

Incorporating the inequity aversion transformations, especially in the framed treatment where participants were aware of their competition, aimed to provide a more nuanced representation of participants' behavior.

3.3 Experimental Design

In our study's experimental phase, we used a 2×2 non-constant sum game with the cyclic game structure. The following sections of this chapter provide more information on this design.

3.3.1 Game Mechanics

Our game modeled the interaction between a student and an examiner during an exam scenario. The student has the strategy options "to cheat" (U) or "not cheat" (D), while the examiner can choose between "control" (L) and "not control" (R). In our 2×2 context, "control" implies that "cheating" will always be detected. We use "¬" to represent the negation-based strategies, respectively ¬Cheat (not to cheat) and ¬Control (not to control). Figure 3.1 shows the applied game matrix.

		STUDENT (PLAYER 2)	
		Cheat (L)	Not Cheat (R)
INSPECTOR (PLAYER 1)	Control (U)	12, 2	11, 10
	Not Control (D)	3, 16	14, 14

Figure 3.1: Inspector Game Matrix

In our design, we decided to allow for the possibility of coordination to give a better chance for the Nash equilibrium, as it could imply more minor deviations from assumptions of strict rationality. In the original experiment by Selten and Chmura (2008), the players were matched randomly within pre-defined groups to avoid the possibility of coordination. In this experiment, the players were matched randomly at the start, and the same pair continued in the same game until the end. Another slight design difference is the number of iterations for each game-play. We observed in preliminary tests that the choice probabilities results became stable much earlier (after around 30 periods) than in the 200 periods proposed originally. Therefore, we reduced this number to 70 to make the experiment more efficient. Because we assume that ”cheating” is not the norm in a real-life situation, ”control” implies effort or costs. The Nash equilibrium has a slight -presumably realistic-tendency for $(\neg Control, \neg Cheat)$. In this case, the preferred outcome for the inspector player is $(\neg Control, \neg Cheat)$, and for the student $(\neg Control, Cheat)$.

3.3.2 Framing Schemes

Information framing can significantly influence decisions, even when the underlying information remains constant. Tversky and Kahneman (1981) highlighted the interplay of cognitive biases in decision-making, a finding also shared by Andreoni (1995); Cookson (2000), who noted how different descriptions of a single choice can lead to varied outcomes. In this study, we introduce both neutral (”context-free”) and information-loaded (”in-context”) frames to gauge their effects on perceptions, particularly in areas with potential ethical implications imposed through cheating and controlling behavior actions.

This design choice mirrors the approach of Abbink and Hennig-Schmidt (2006), which examined bribery through a similar lens. While there is a mix of findings regarding the impact of such frames—Abbink and Hennig-Schmidt (2006); Abbink, Irlenbusch, and Renner (2002) reported negligible effects, whereas Dufwenberg, Gächter, and Hennig-Schmidt (2011) noted distinct behavioral shifts using neutral frames—we aim to look further into potential variances referencing our methods as the ”framed” and ”unframed” treatments throughout our research. See appendix 3.6.4 for details about the experiment design and frames.

3.4 Results

This section presents the results of the inspector game experiment. We begin with general statistics and then apply stationary equilibrium concepts across aggregated data, time series, and individual games. We also apply the modified concepts to other experiments. Our accuracy assessment method, based on Selten and Chmura (2008), measures the distance between predictions and observed outcomes, denominated as Q (details in appendix 3.6.2).

In our sample, we had 100 student participants [1] in laboratory sessions: 52 in the framed game (26 observations) and 48 in the unframed game (24 observations). The group was

[1] Note: here, we have to distinguish the student participants from the role of ”student” in the game. All participants were students. However, in the game, they were randomly assigned either ”student” or ”inspector” roles.

58% female and 42% male, primarily from Rhine-Waal University in Kleve, Germany. Half were German, while the rest were diverse, including 7% Russian and 6% Indian. The average completion time was 40 minutes, with participants earning 16.75 euros on average.

3.4.1 Game Behavior and Framing Analysis

We compared data samples between different groups, including framed and unframed treatments, male and female gender, and each player type individually, using the frequency distributions of individual player strategies aggregated over 70 periods. Non-significant treatment differences were found using both a non-parametric Mann–Whitney U test (Mann & Whitney, 1947) and a permutation test (Fisher, 1936). The complete overview of the sample comparison results is documented in table 3.4, appendix 3.6.1. These results align with Abbink and Hennig-Schmidt (2006); Abbink et al. (2002), where neutral frames did not significantly affect participants' perceptions of the task. Figure 3.2 describes the overall game behavior analysis.

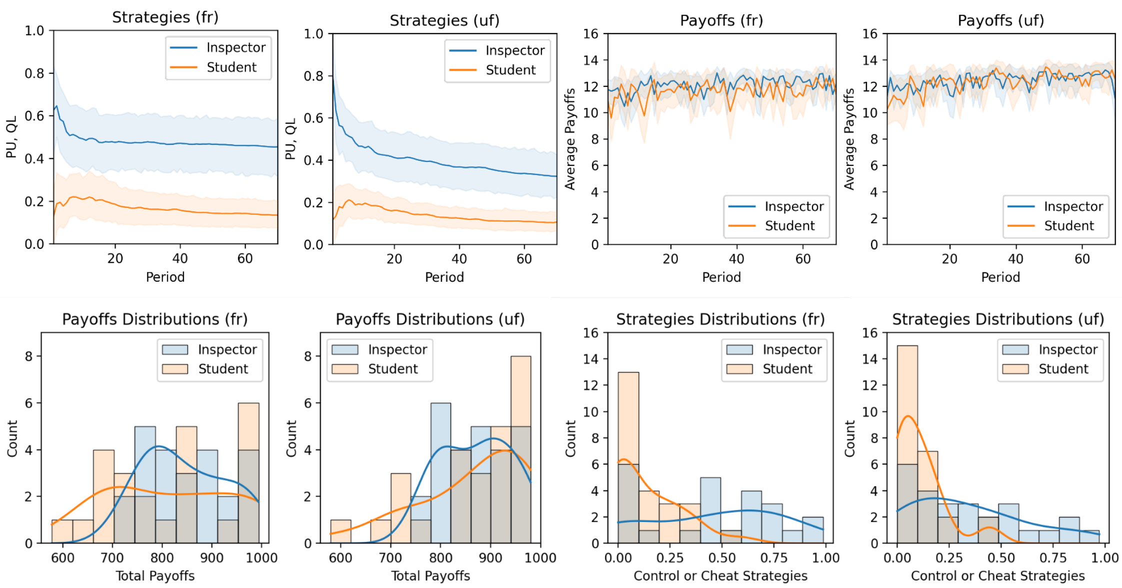


Figure 3.2: Game Behavior Analysis. From left to right: in the first row, we have the subjects' observed strategy selection behavior and payoffs per round. On the second row, we have the distributions of total payoffs and the distributions of cheat and control frequencies .

Interestingly, despite similar strategy frequency distributions, both examiners and students played the strategies control (strategy U, +12%) and cheat (strategy L, +3%) more often in the framed games, which are strategies associated with higher risks. This effect was reflected in the payoffs at the end of the experiment. Examiners scored, on average, 846 points (12.1 per round) in framed games versus 869 points (12.4 per round) in unframed games. Students' average score in framed games was 814 (11.6 per round), while in unframed games, it was 863 (12.3 per round). This suggests subjects engaged in riskier behavior when they had complete information about the game, resulting in slightly lower payoffs in the framed version. These differences are visible in figure 3.2, in both strategy selection behavior and payoff results.

Brandts and Schwieren (2007, 2009) showed that different framing designs and incentives lead to diverse effects, with more intense interventions producing larger effects. Our study used a neutral vs. non-neutral framing scheme, which may not have been strong enough to generate significant treatment effects in this game. Individual decision-making frequencies yielded similar results for most analyzed features.

3.4.2 Aggregated data

In the aggregated data level, we calculated average choice probabilities for each player type in both treatments, bundled across subjects and iterations. The distribution of choice probabilities, as shown in Figure 3.3, reveals that the examiner profile (p_U) is more scattered than the student profile (q_L), which is more concentrated. The distribution information for the examiner profile includes a variance of 0.10 (framed) and 0.07 (unframed) and a standard deviation of 0.32 (framed) and 0.27 (unframed). For the student profile, we observed a variance of 0.03 (framed) and 0.01 (unframed) and a standard deviation of 0.16 (framed) and 0.12 (unframed). Each point in Figure 3.3 represents a mixed strategy profile (p_U, q_L) for one observation, and the cross indicates the means. For the aggregated data, mean frequencies were $(p_U = 0.461, q_L = 0.136)$ for the framed treatment and $(p_U = 0.341, q_L = 0.108)$ for the unframed treatment.

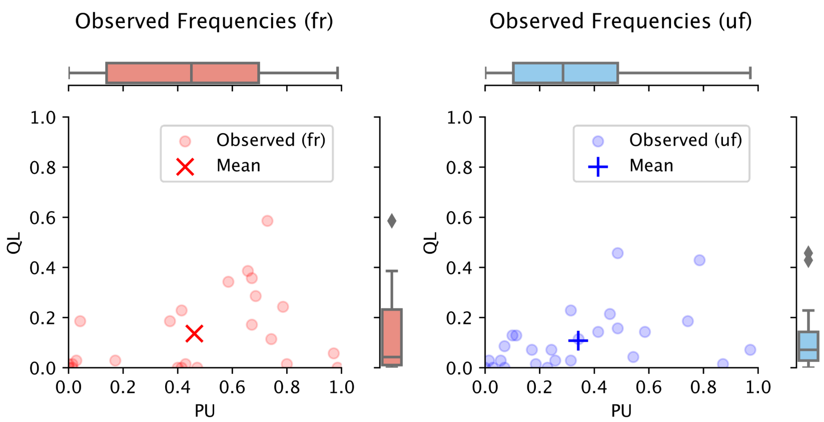


Figure 3.3: Observed relative frequencies distributions for the framed and unframed treatments

We compared our empirical results to a total of ten stationary equilibrium concepts' predictions. Table 3.1 provides a complete overview of the equilibria computation results. The table also depicts the parametric models' selected parameter values and the distances (Q) to the empirical observations. Except for the NE, most predictions yield satisfactory results. The IBE and ASE models are the best predictors among the standard concepts. One can also observe that the modified concepts generally seem to fit the data better. This could imply that parametric modifications yielded more flexible or adaptive models in capturing the complexities of empirical data.

	Framed Treatment				Unframed Treatment			
Standard Concepts	PU	QL	Q	Params.	PU	QL	Q	Params.
NE	0.200	0.250	0.285	-	0.200	0.250	0.200	-
IBE	0.380	0.170	0.088	-	0.380	0.170	0.073	-
QRE	0.426	0.208	0.080	$\lambda = 0.591$	0.342	0.195	0.087	$\lambda = 0.995$
ASE	0.440	0.176	0.045	$N = 3$	0.349	0.180	0.072	$N = 4$
PSE	0.431	0.203	0.073	$N = 7$	0.382	0.203	0.103	$N = 11$
Modified Concepts	PU	QL	Q	Params.	PU	QL	Q	Params.
$IBE\gamma$	0.423	0.196	0.071	$\gamma = 1.4$	0.320	0.135	0.034	$\gamma = 3.2$
$NE \circ IA_{(\alpha,\beta)}$	0.200	0.250	0.285	$\alpha = 0$ $\beta = 0$	0.180	0.220	0.196	$\alpha = 0.127$ $\beta = 0$
$IBE \circ IA_{(\alpha,\beta)}$	0.461	0.136	0.000	$\alpha = 0.768$ $\beta = 0.172$	0.374	0.165	0.065	$\alpha = 1$ $\beta = 0.991$
$IBE_2 \circ IA_{(\alpha,\beta)}$	0.461	0.136	0.000	$\alpha = 0.763$ $\beta = 0.847$ $\gamma = 2$	0.341	0.108	0.000	$\alpha = 0.886$ $\beta = 0.543$ $\gamma = 2$
$IBE\gamma \circ IA_{(\alpha,\beta)}$	0.461	0.136	0.000	$\alpha = 0.617$ $\beta = 0.146$ $\gamma = 1.1$	0.341	0.108	0.000	$\alpha = 0.069$ $\beta = 0.578$ $\gamma = 3.7$

Table 3.1: Compiled results of the application of all stationary concepts

The IBE, a model rooted in behavioral impulses, demonstrates notable disparities in predictions across the same game framed differently. Such framing likely heightens specific behavioral tendencies, potentially diminishing participants' loss aversion. In comparing it with the parametric version, adjusting the loss aversion values through the γ parameter brings the model's predictions closer to empirical findings, accounting for the framing effects.

This predictive power comparison is clearly outlined in figure 3.4. The plots rank the Q values from highest (worst) to lowest (best), showing the standard concepts first, followed by the modified concepts. Overall, the ASE and IBE concepts were the best predictors for our treatments, while NE was the least effective. QRE performed better in the unframed treatment, ranking second, but less effective in the framed treatment, being the second-worst predictor. PSE ranked third in both treatments.

For the modified concepts, the IBE model variants provided the best predictions, with a closer fit to the empirical data (see table 3.1). $IBE_\gamma \circ IA_{(\alpha,\beta)}$ had the best overall predictive power in both treatments. The other IBE modifications, $IBE \circ IA_{(\alpha,\beta)}$ and $IBE_2 \circ IA_{(\alpha,\beta)}$ similarly demonstrated satisfactory accuracy, particularly when compared to the original concepts. IBE_γ exhibited superior performance compared to the fixed-parameter version.

Selten and Chmura (2008)'s results [2] rendered the NE as the worst performer, reinforcing that Nash Equilibrium is a poor predictor for games with only mixed strategies. The authors ranked concept performance as 1. IBE, 2. PSE, 3. ASE, 4. QRE, and 5. NE.

[2]Note: Unlike the authors, who used 12 distinct games and averaged parameter values for the QRE, ASE, and PSE models, we estimated parameters for a single game type. This approach removed the need for a sampling variance component in our Q values.

Our rankings share similarities; however, IBE performs well in the framed version of our experiment but not as well in the unframed version, being superior only to NE predictions.

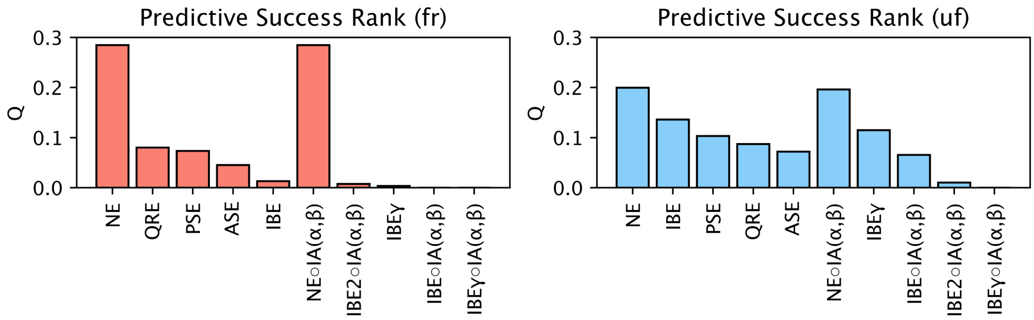


Figure 3.4: Predictive performance measure Q, ranked for the standard and modified concepts

Among the modified concepts, the NE modification, $NE \circ IA_{(\alpha,\beta)}$, had nearly the same prediction accuracy as the original NE. For the framed treatment, it remained the same, considering up to 3 decimal points, while in the unframed treatment, there was a slight change, resulting in a reduction of 0.004 in the Q value. Generally, incorporating behavioral parameters greatly enhances the accuracy of the IBE compared to the NE. The rigid rationality assumption of NE, which prioritizes individual payoff maximization, might limit its adaptability to behavioral nuances. In contrast, IBE, rooted in behavioral impulses, is inherently flexible and captures the complexities of real-world decision-making. Consequently, when behavioral elements are introduced, IBE naturally aligns better with observed human behavior, while NE's foundational assumptions may overshadow these adjustments.

Proceeding with the empirical fit analysis, figure 3.5 displays the prediction values from our analysis, giving an overview of the concepts' fit to empirical observations. The plots are divided into treatments and standard versus modified concepts. Empirical observations are shown in red (framed treatment) and blue (unframed treatment), and different marker colors and shapes represent equilibrium concepts. Except for NE, standard concepts are closely grouped with minor Q value differences. Modified concepts show a better overall fit. After incorporating the IA parameters in NE, predicted values showed minimal differences, indicating that no parameter value in the search space could produce better predictions.

We sought to understand how the parameter values from the parametric models affect the predicted probabilities. For a visual exploration of the QRE λ and the introduced γ for the IBE, figure 3.6 displays the development of the parameters during the optimization process. QRE performed as expected; there was a convergence for higher lambdas to the Nash equilibrium. The closest (p_U, q_L) prediction values to the empirical observations are given as $(0.426, 0.208)$ with $\lambda = 0.591$ for the framed treatment and $(0.342, 0.195)$ with $\lambda = 0.995$ for the unframed treatment. The QRE search space represented by this curve starts at a given pair of predicted probabilities with low λ values, and as the quantal response functions become steep, it gets closer to the NE.

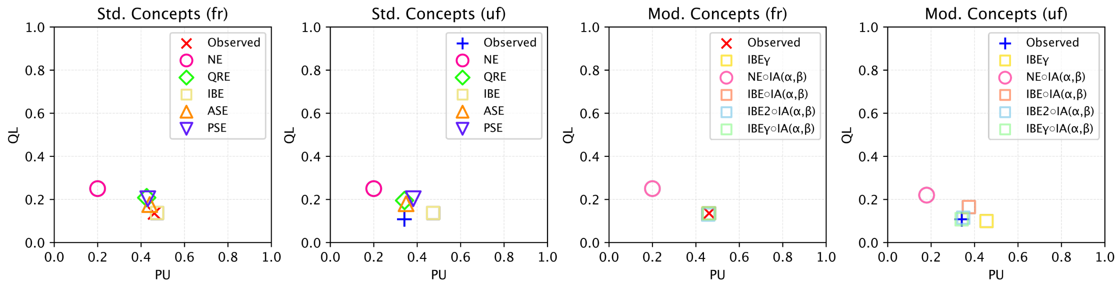


Figure 3.5: Predictions for the standard stationary equilibrium concepts c to the aggregated observed relative frequencies

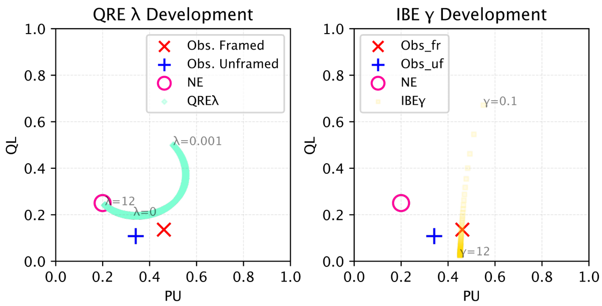


Figure 3.6: Development of λ and γ parameters in the search space for the best predictions

For the IBE moodels, the closest predictions are $(0.457, 0.135)$ with $\gamma = 1.9$ for the framed treatment and $(0.455, 0.100)$ with $\gamma = 2.7$ for the unframed treatment. As with the QRE and IBE models, the γ parameter increases as it progresses through the search space until it reaches a limit. The results show that the framed treatment has a lower overall loss aversion (γ) value than the unframed treatment, consistent with the statistical tests showing that subjects with more information tended to play riskier strategies. This finding suggests that the γ parameter can represent the loss aversion behavioral trait in the IBE-based models.

As the second behavioral measurement, we assessed inequity aversion by examining strategy changes after inequitable outcomes, with stronger aversion expected in the unframed treatment where subjects were unaware of playing against each other. We calculated conditional frequencies of strategy changes in response to the worst outcomes, determining inequity aversion by the frequency of these changes relative to maximum payoff differences. Envy indicated strategy changes after a player's losses, while guilt corresponded to changes after an opponent's loss. For envy, the results were framed $= 0.74$, unframed $= 0.57$; for guilt, framed $= 0.64$ and unframed $= 0.67$. The means of α and β across all IA-based modes are framed $\alpha = 0.35$, $\beta = 0.08$ and unframed $\alpha = 0.48$, $\beta = 0.64$. There is some consistency in comparing the results in the unframed but not in the framed treatment. Measuring this is challenging due to potential strategy changes caused by the game's asymmetric payoff differences, which depend on the matrix structure since it also showed strong values in the framed treatment. The detailed analysis is in appendix 3.6.1.

A more in-depth analysis may be necessary to understand the relationship between these parameters and actual behavior better.

For the sampling-based equilibria methods ASE and PSE, we checked the sample sizes $N = 1, ..., 12$. The best fit to our data was given by four different sample size N values, and we did not identify an optimal sample size that fits all values well. For ASE, we found $n = 3$ to be the best fit in the framed treatment and $n = 4$ in the unframed treatment. For PSE, we found higher values to be the best fit, with $n = 7$ in the framed treatment and $n = 11$ in the unframed treatment. Figure 3.7 shows the predicted probabilities yielded by both ASE and PSE concepts and the distances Q to the observed relative frequencies. The search space of the sampling equilibria was restricted to the number of sample sizes considered.

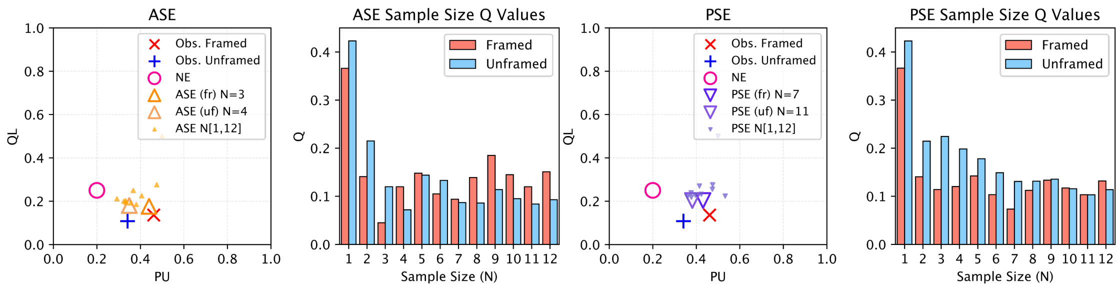


Figure 3.7: ASE and PSE predictions for sample sizes 1-12

We wanted to display the particular search space for all the explored models and how the possible predictions are distributed within the space of choices. Thus, we collected all prediction values yielded by each model in the following parameter ranges specified in appendix 3.6.3. The predictions distribution within the search spaces are plotted in figure 3.8; the heat maps represent the density distribution of all yielded predictions of probabilities for the players' mixed strategies (p_U, q_L).

The superior performance of modified concepts can be attributed to their ability to explore larger, more comprehensive search spaces, as seen in figure 3.8. Their primary advantage lies in exploring larger probability spaces constrained by subjects' behavioral parameters, resulting in these observed high-density predictions close to empirical values. However, search spaces are game-specific and vary according to the game's structure. For example, models with fundamentally smaller search spaces than the three-parameter model also yield accurate predictions, reinforcing their predictive power claims.

3.4.3 Time-Series Analysis

In this section, we enhance the aggregated relative frequencies analysis by adding a time dimension to the study. We compute the equilibrium probabilities considering the cumulative development of strategies chosen by each group of players per period. Additionally, we examine the players' strategic behavior over time by generating stationary equilibrium predictions in each round. The results of this analysis are concisely presented in Figure

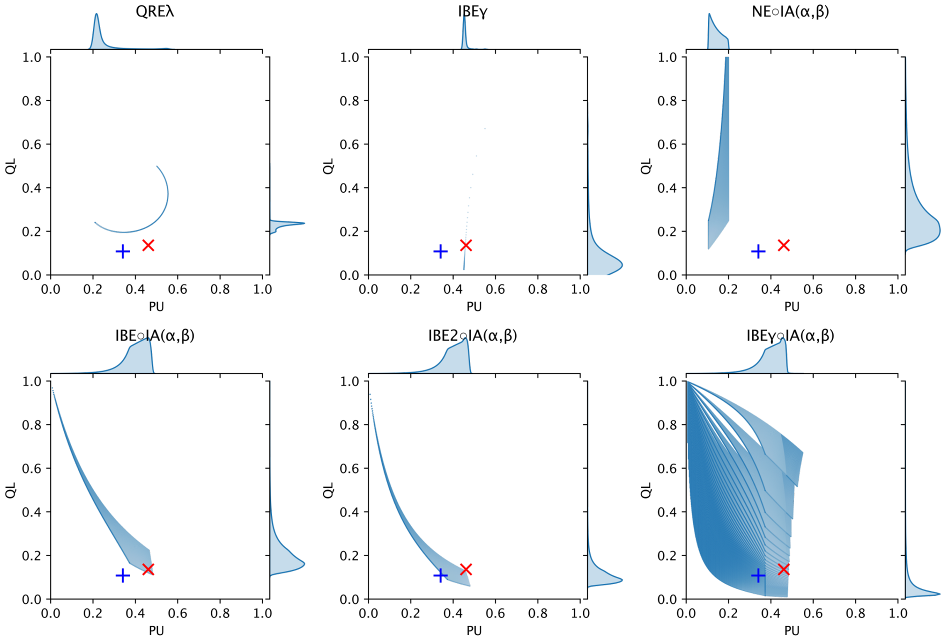


Figure 3.8: Heatmaps with the search space for the parametric models

3.9, which offers a deeper understanding of the evolution of players' strategic behavior throughout the game.

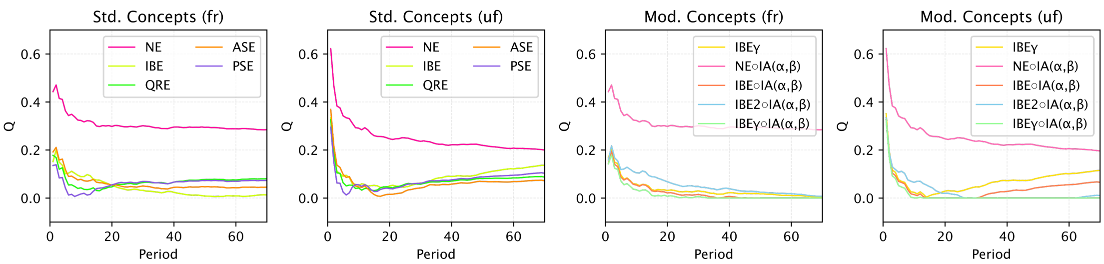


Figure 3.9: Distances Q to observed relative frequencies per period

We observed consistent learning effects in both framed and unframed treatments. Initially, distances Q to model predictions were higher but decreased over time as players gained context understanding and strategy selection improved. Choices converged closer to equilibrium predictions and stabilized after 20 periods. Choi, Gale, and Kariv (2012) found similar results using QRE, where agents learned by observing others' behavior in limited information scenarios. Incorporating natural noise in decision processes enables theoretical models to provide structural decision estimations and yield predictions close to empirical data.

Learning effects were present in varying degrees for all stationary concepts in our study, with models incorporating behavioral traits generally outperforming strictly rational predictions. The effects are evident in the empirical choice probabilities when the distribution is split before and after the 20th and 40th periods (table 3.2).

	Framed Treatment		Unframed Treatment	
Period split	p_U	q_L	p_U	q_L
1-20	0.492	0.188	0.431	0.163
21-40	0.471	0.131	0.338	0.090
41-70	0.433	0.104	0.282	0.083

Table 3.2: Relative frequencies of the empirical observations with period split

Both treatments showed the highest differences between the first 20 and the remaining periods. However, we observed that the framing affected the learning process, as the participants of the framed experiment presented lower Q values in the first periods and displayed a more stable strategy selection behavior if compared to the participants in the unframed treatment. We have also isolated the individual parameter values in applying parametric models on a period level. The compiled results are presented in figure 3.10. Each chart represents the parameter value for each round's predictions for the equilibrium concept described in the title of the charts.

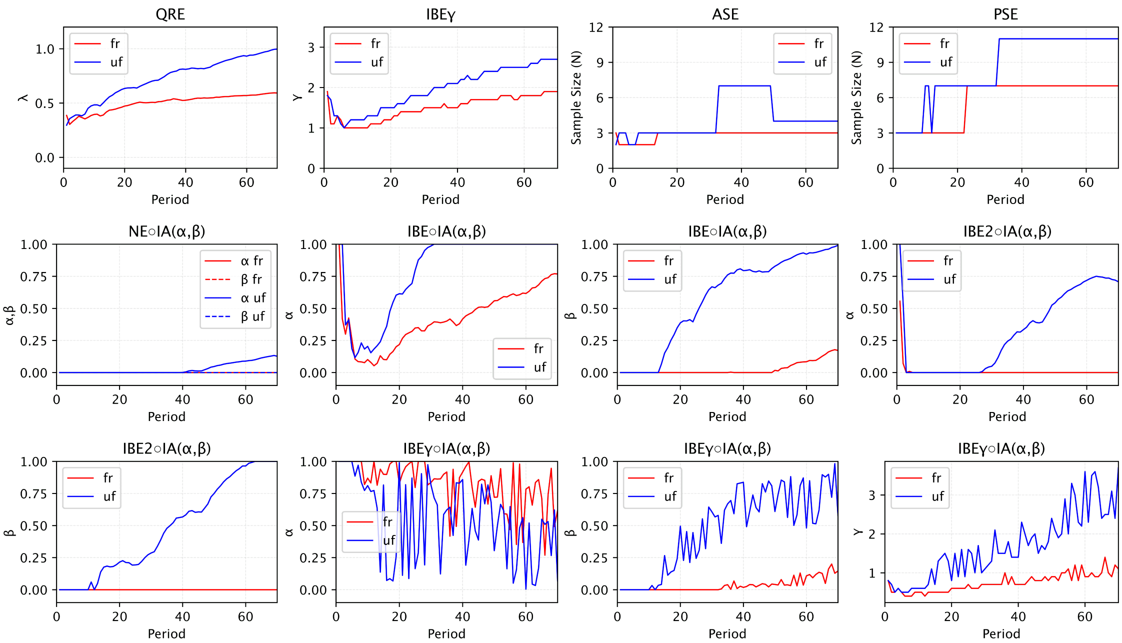


Figure 3.10: Parameters development per period, for all parametric models

For a better understanding of the parameters' meaning, we assigned them to their respective theoretical explanations in each of the models:

- QRE - λ: precision of beliefs about others' strategies (McKelvey & Palfrey, 1995)
- IBE - γ: loss aversion (Ockenfels & Selten, 2005; Selten & Chmura, 2008)
- ASE - N: memory about other players' past actions (Osborne & Rubinstein, 2003; Selten & Chmura, 2008)
- PSE N: memory about own past payoffs (Osborne & Rubinstein, 1998; Selten &

Chmura, 2008)

- IA α: envy, β: guilt (Fehr, Naef, & Schmidt, 2006; Fehr & Schmidt, 1999; Rey-Biel, 2008)

This explanation assists in interpreting the charts in Figure 3.10. We observed a near-monotonic increase over time for the bounded rationality trait (QRE λ), indicating that predictions moved closer to the NE and became more constant in later rounds. A similar trend was seen for the loss aversion trait (IBE γ); subjects initially displayed riskier behavior, progressing to a more loss-averse state over time. Referring to Table 3.2, higher incidences of "cheat" and "control" strategies were observed early in the experiment, reflecting riskier behavior consistent with theoretical assumptions.

As memory improves in sampling equilibria, an ideal sample size is found and only modified if another sample size proves superior. We observed more stable values for both cases (ASE N and PSE N), with few changes across the 70 periods. Trends for the inequity aversion's envy and guilt values (α and β) were less clear. While an increasing trend was observed most of the time, more significant variations occurred from round to round, especially in the 3-parameter model $IBE\gamma \circ IA_{(\alpha,\beta)}$. In the unframed treatment, the values of the IA parameters were significantly higher and increased faster than in the framed treatment. The behavioral interpretation suggests that inequity aversion had a more substantial influence on decisions in the unframed treatments. Except for the α parameter in the $IBE\gamma \circ IA_{(\alpha,\beta)}$ model, the development of the parameters remained consistent across models and treatments.

3.4.4 Game-Level Analysis

In the final section of our results analysis, we examined observations on a game-play level. To do this, we considered each pair of players' strategy choices over 70 periods as one observation, aggregating pairs of probabilities (p_U, q_L) for students and inspectors. We analyzed 24 gameplays in the framed treatment and 26 in the unframed treatment, individually computing the Q values for all stationary concepts. Table 3.3 displays the frequency with which each concept was ranked against the others, based on prediction power for each observation. In other words, it shows how often each concept was ranked between 1 (best) and 10 (worst) based on the lowest Q values.

We observed consistency with the aggregated analysis, highlighting a higher predictive power for the parametric models. The modified concepts extensively occupied the ranks' best positions, with $IBE\gamma \circ IA_{(\alpha,\beta)}$ demonstrating the best performance most of the time. The other concepts were distributed more evenly across the ranks. Interestingly, the modified NE version $NE \circ IA_{(\alpha,\beta)}$, QRE, and PSE also exhibited significant performance in the first position frequency despite not performing as well on the aggregated level. In contrast, the original IBE did not fare as well in the observation-level ranks, while the $IBE\gamma$ showcased a significantly superior performance. For the IBE case, this isolated game analysis using a dynamic parameter dramatically improved the accuracy of the predictions.

Fr. Treatment Rank	1	2	3	4	5	6	7	8	9	10
NE	0%	0%	0%	4%	19%	4%	4%	25%	50%	4%
IBE	0%	0%	0%	0%	0%	0%	4%	0%	31%	75%
QRE	0%	0%	25%	19%	19%	29%	8%	0%	0%	0%
ASE	0%	0%	0%	22%	19%	13%	42%	3%	0%	0%
PSE	0%	0%	13%	30%	5%	8%	13%	0%	13%	21%
$IBE\gamma$	21%	29%	8%	22%	10%	4%	4%	0%	0%	0%
$NE \circ IA_{(\alpha,\beta)}$	0%	4%	25%	0%	0%	0%	0%	53%	0%	0%
$IBE \circ IA_{(\alpha,\beta)}$	0%	8%	21%	0%	14%	8%	21%	19%	6%	0%
$IBE_2 \circ IA_{(\alpha,\beta)}$	0%	38%	8%	4%	14%	33%	4%	0%	0%	0%
$IBE\gamma \circ IA_{(\alpha,\beta)}$	79%	21%	0%	0%	0%	0%	0%	0%	0%	0%
Uf. Treatment Rank	1	2	3	4	5	6	7	8	9	10
NE	0%	0%	0%	0%	20%	15%	4%	31%	24%	4%
IBE	0%	0%	0%	0%	0%	0%	0%	0%	35%	77%
QRE	0%	4%	4%	23%	44%	23%	4%	0%	0%	0%
ASE	0%	0%	4%	15%	12%	15%	54%	0%	0%	0%
PSE	0%	0%	11%	12%	4%	8%	15%	3%	41%	19%
$IBE\gamma$	20%	23%	26%	8%	16%	4%	4%	0%	0%	0%
$NE \circ IA_{(\alpha,\beta)}$	4%	8%	26%	8%	4%	0%	0%	37%	0%	0%
$IBE \circ IA_{(\alpha,\beta)}$	0%	8%	11%	15%	0%	12%	15%	29%	0%	0%
$IBE_2 \circ IA_{(\alpha,\beta)}$	0%	35%	19%	19%	0%	23%	4%	0%	0%	0%
$IBE\gamma \circ IA_{(\alpha,\beta)}$	76%	23%	0%	0%	0%	0%	0%	0%	0%	0%

Table 3.3: Ranked stationary concept frequency in terms of predictive power for each model applied to all individual observations, on a game-play level

3.4.5 Modified Equilibria Concepts Applied to Other Experiments

To evaluate the predictive power of the modified concepts introduced earlier, we applied them to the original 12 games experiment introduced by Selten and Chmura (2008) and later adjusted by Brunner et al. (2010). The consolidated results for each combination of game and concept, including the outcomes for p_U, q_L, and Q, as well as the fitted parameters for each combination, are documented in table 3.6, appendix 3.6.1. The results exhibit a similar pattern to those of our game. The modified concepts generally demonstrate improvements when compared to their original versions. In terms of predictive performance, based on our Q measure, we observe a ranking akin to the one presented in Figure 3.4. Considering the mean Q values for all 12 games, we obtain the following order from most to least accurate predictor: $IBE_\gamma \circ IA_{\alpha,\beta}$ (mean $Q = 0.002$), $IBE \circ IA_{\alpha,\beta}$ (mean $Q = 0.007$), $IBE_2 \circ IA_{\alpha,\beta}$ (mean $Q = 0.032$), IBE_γ (mean $Q = 0.038$), and $NE \circ IA_{\alpha,\beta}$ (mean $Q = 0.131$).

The predictive performance for each game and equilibrium concept is shown in Figure 3.11. The 3D bar plot displays the 12 games on the X-axis, the categorical equilibrium concept on the Y-axis, and the Q values on the Z-axis, represented by the bars' size. The concepts are ordered by the mean Q values for all games, from lowest to highest. Note that our Q formula (defined in appendix 3.6.2) slightly differs from the one in Selten and Chmura (2008). We use the Euclidean distance formula, while they employ a quadratic distance formula and account for sampling variance since it is a multi-game experiment.

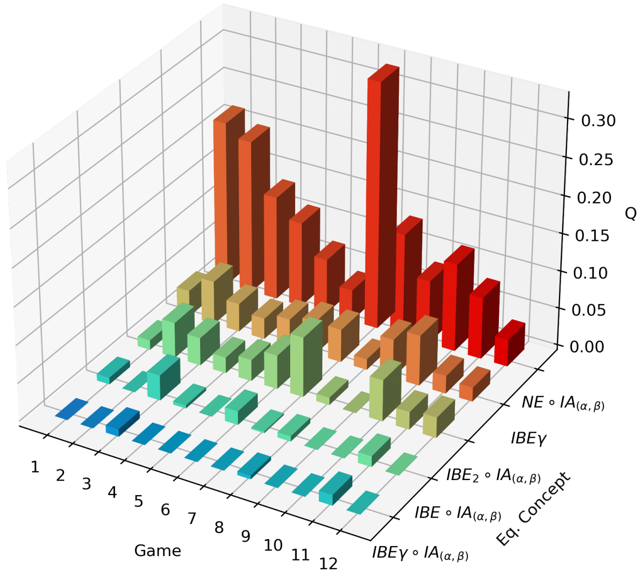


Figure 3.11: Measure of predictive success Q for the modified concepts applied to the 12 games

This comparison reveals that the search functions generated by the modified concepts are game-specific. Similar to the plots in Figure 3.8, we plotted the search space functions for all 12 games in Figures 3.12 and 3.13 (appendix 3.6.1), with "X" indicating the observed frequencies. The IBE and NE variations are game-specific, with different search functions depending on the game structure. Models with more parameters generate more prediction possibilities. However, in every case, the models' predictions are relatively close to the empirical values, reinforcing their flexibility and adaptability to different game scenarios.

3.5 Conclusions and Discussion

This paper extends the stationary equilibrium literature, building on Selten and Chmura (2008) by employing a different game, applying different frames, and allowing coordination. We introduced Loss Aversion and Inequity Aversion conceptual modifications and analyzed aggregated data, time series, and game-level observations. Our study offers another perspective on the stationary equilibria theory. The Inspector game framework was interesting in this context as it introduces several behavioral-related mechanisms connected to both the incentive schemes in the game and the contextual and personal influences regarding cheating and punishment actions.

Our game design represents the idea that cheating is less likely than not cheating, and controlling is less likely than not controlling. Additionally, the probabilities of the NE represent a more straightforward pattern by using inverse "prominent " numbers - 0.20 and 0.25 for the first strategy of players 1 and 2, respectively - than the probabilities of the games reported in Selten and Chmura (2008). Though we allowed additional possible coordination, this did not lead to a better prediction of the empirical data by the NE. We showed that in a situation with unclear and vague norms like a mixed NE, players do not recognize the norm but try instead to follow an impulse to reduce losses.

The information-loaded and neutral frames had a small and statistically insignificant effect, where participants with complete information played the riskier strategies more often. In the computation of stationary predictions, we ranked the applied concepts as follows: framed treatment: 1. IBE, 2. ASE, 3. PSE, 4. QRE, 5. NE; unframed treatment: 1. ASE, 2. QRE, 3. PSE, 4. IBE, 5. NE. In line with the literature, stationary concepts outperformed. Four of our concept modifications demonstrated improved predictive power, except for the NE modification, which showed better results only at the observation level. While models with additional parameters enhance accuracy, they also increase computational complexity and pose interpretational challenges.

Our parameter investigation outlined that theoretically quantifying behavioral parameters does a fair job of capturing real-world behavior. Our IBE parameter for loss aversion showed lower loss aversion values for riskier players and greater loss aversion values for players who choose high-risk strategies less frequently. That fact was also outlined by our framing intervention, where players with more information about the context appeared to take more risks. The time-series analysis also reinforced this finding, displaying learning effects and equilibrium stability. Subjects selected riskier strategies in the first 20 periods. The strategic behavior converged closer to the theory predictions over time after learning about the environment, and their loss aversion measurements rose over time, analogous to equilibrium convergence. Regarding inequity aversion, we found no substantial evidence that envy and guilt matched the empirical behavior, even if they made sense according to the situations applied.

Our study's limitations include its focus on a specific game, potentially limiting generalizability to other scenarios. We lean on behavioral traits and information constraints to explain decisions, which might be a simplified perspective on the intricate nature of human choices. While laboratory settings provide controlled environments essential for initial observations on decision-making, real-world data, with its inherent complexity, offers a deeper and more robust understanding of mixed-strategy behaviors, as highlighted by Mauersberger and Nagel (2018). Furthermore, introducing new parameters poses interpretational challenges for practitioners, increasing the calculation complexity.

In summary, our paper builds upon the literature groundwork in stationary equilibria by introducing modified versions of these concepts, exploring the influence of behavioral traits and framing effects, and providing a deeper analysis of model parameters and their temporal dynamics. The main contribution of this paper is a more comprehensive understanding of equilibria in completely mixed games.

3.6 Appendix

3.6.1 Additional Analyses

Table 3.4 shows the p-values for the non-parametric tests employed to compare the data samples of interest.

Table 3.5 displays the results of the inequity aversion analysis. Here, we assess inequity aversion in both framed and unframed treatments by examining strategy changes post-

Sample A	Sample B	Test Results
Framed	Unframed	Mann-Whitney U: 0.614, Permutation: 0.092
Male	Female	Mann-Whitney U: 0.904, Permutation: 0.291
P1 Framed	P1 Unframed	Mann-Whitney U: 0.256, Permutation: 0.083
P2 Framed	P2 Unframed	Mann-Whitney U: 0.946, Permutation: 0.253

Table 3.4: Strategy Selection Statistical Tests Overview - p Values

inequitable outcomes. We have employed two forms of aversion computation based on our behavioral parameters - envy (strategy changes after a player's losses) and guilt (changes after an opponent's loss).

	Unframed		Framed	
	Examiner	Student	Examiner	Student
Strategy Change Envy	0.29	0.84	0.60	0.89
Strategy Change Guilt	0.53	0.81	0.42	0.86

Table 3.5: Analysis of strategy changes based on Inequity Aversion

In the presented data, players exhibited a pronounced inclination towards envy-driven strategy changes in the framed context, with a change rate of 0.74 compared to 0.57 in the unframed scenario. Conversely, guilt-driven strategic alterations were slightly more prevalent in the unframed environment, registering at 0.67, compared to 0.64 in the framed setting (both comparisons use average values). However, there's a notable inconsistency when we examine the IA-based models: while the envy parameter (α) in the unframed setup aligns with observed behaviors, rising to 0.48, the substantial surge in the guilt parameter (β) to 0.64 doesn't mirror the empirical strategy change data.

The remainder of this chapter focuses on the analysis of the equilibria in the 12 additional games. Similar to figure 3.8 in chapter 3.4.2, we have computed all possible prediction points for each of the 12 games from Selten and Chmura (2008). The calculation methods and parameter ranges were kept exactly the same as those used in the inspector game in our experiment. The plots are divided into two figures, figure 3.12 containing the plots for games 1 to 6, and figure 3.13 containing the plots for games 7 to 12. Lastly, table 3.6 displays the details regarding the equilibria computation for each game, including the predicted probabilities (p_U, q_L), the measure of fit Q, and the fitted parameter values.

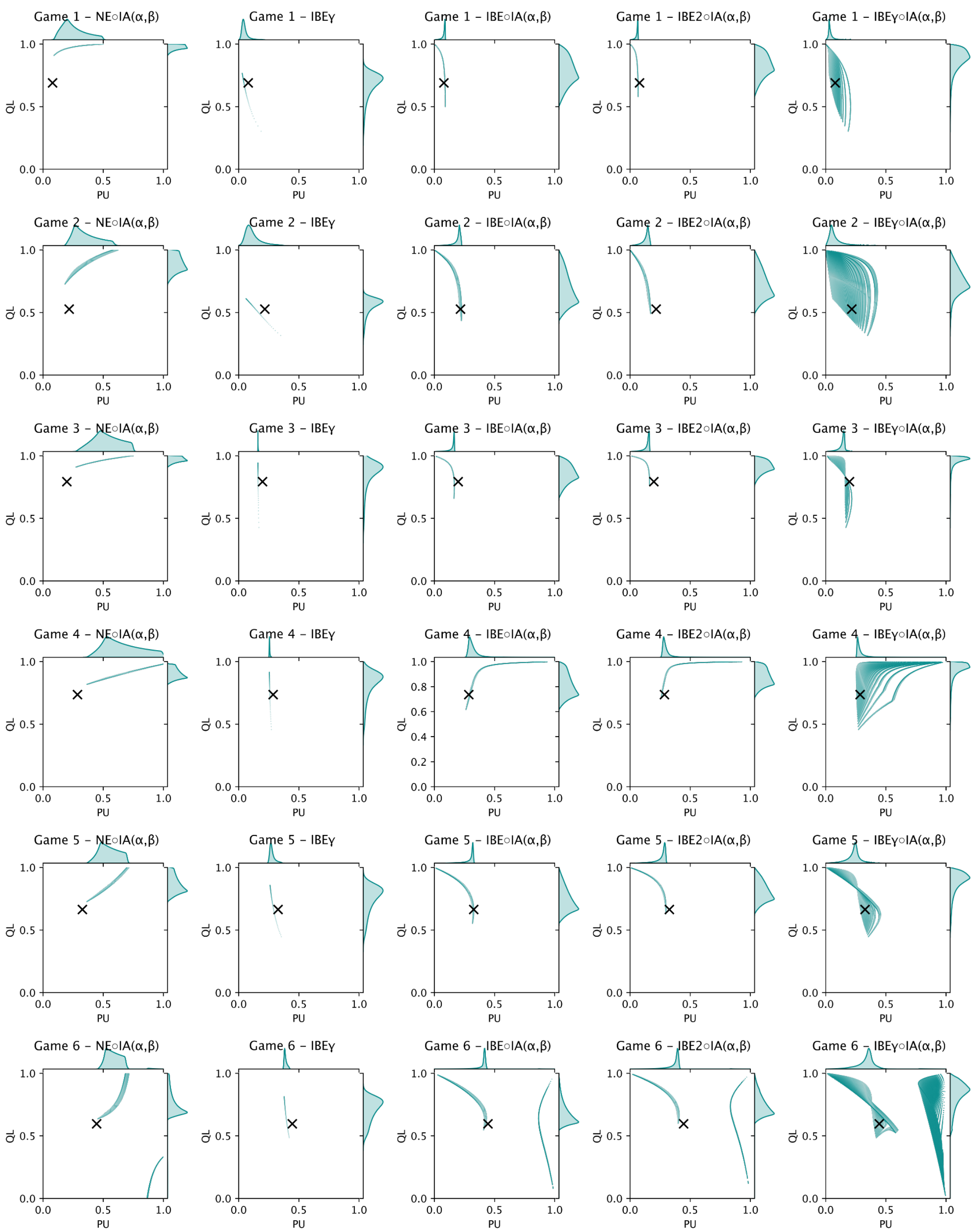


Figure 3.12: Modified Stationary Equilibria Concepts Applied to the games from Selten and Chmura (2008), games 1-6

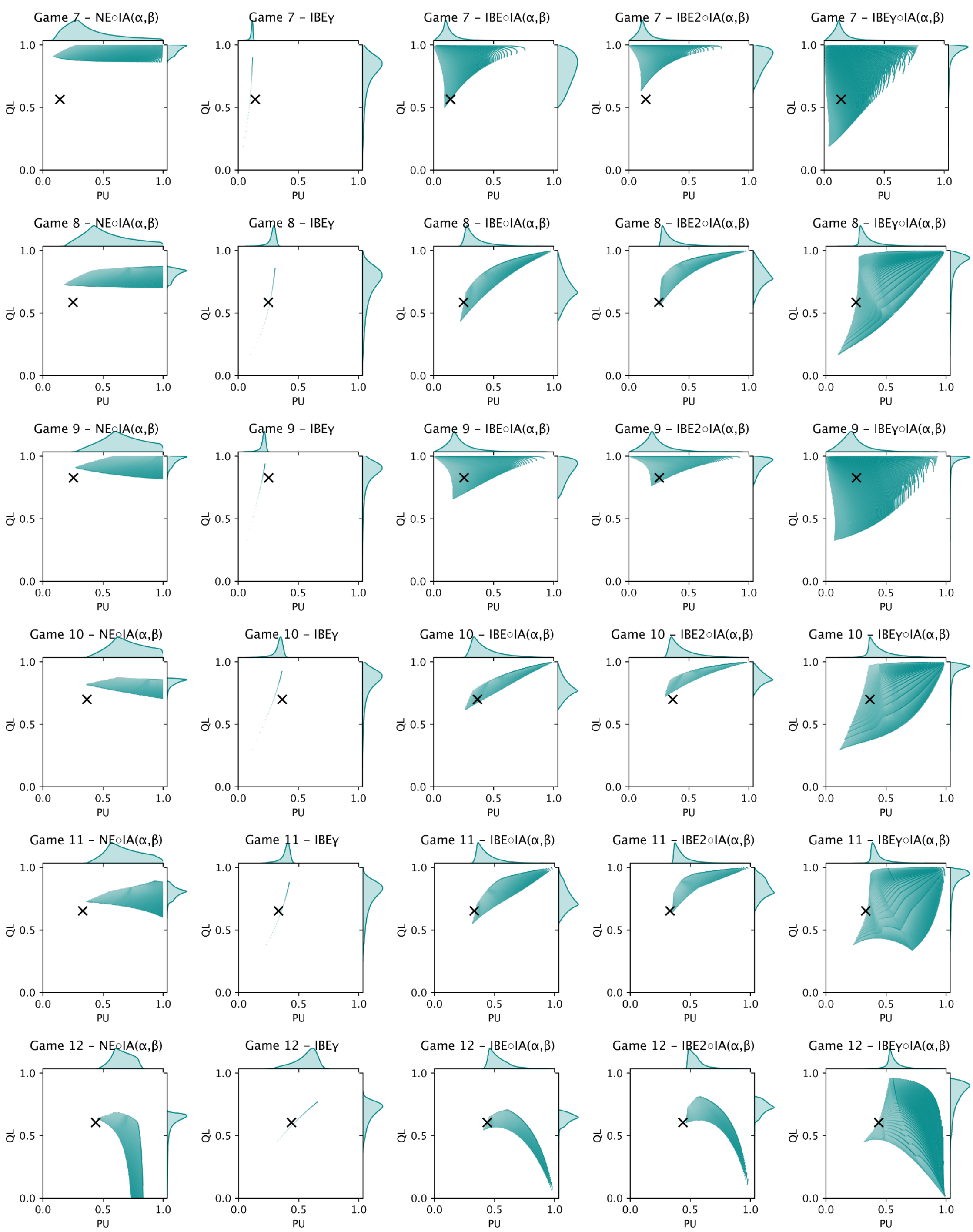


Figure 3.13: Modified Stationary Equilibria Concepts Applied to the games from Selten and Chmura (2008), games 7-12

	Observed		$IBE\gamma$				$NE \circ IA_{(\alpha,\beta)}$				$IBE \circ IA_{(\alpha,\beta)}$				$IBE_2 \circ IA_{(\alpha,\beta)}$				$IBE\gamma \circ IA_{(\alpha,\beta)}$			
Game	p_U	q_L	P.	p_U	q_L	Q	P.	p_U	q_L	Q	P.	p_U	q_L	Q	P.	p_U	q_L	Q	P.	p_U	q_L	Q
1	0.079	0.69	$\gamma = 5.1$	0.044	0.684	0.035	$\alpha = 0$ $\beta = 0$	0.091	0.909	0.219	$\alpha = 0.718$ $\beta = 0.005$	0.088	0.690	0.009	$\alpha = 0.223$ $\beta = 0.071$ $\gamma = 2$	0.066	0.690	0.013	$\alpha = 0.262$ $\beta = 0.142$ $\gamma = 1.3$	0.079	0.690	0.000
2	0.217	0.527	$\gamma = 1.9$	0.175	0.487	0.058	$\alpha = 0$ $\beta = 0$	0.182	0.727	0.203	$\alpha = 0.265$ $\beta = 0.098$	0.217	0.527	0.000	$\alpha = 0$ $\beta = 0.09$ $\gamma = 2$	0.170	0.525	0.047	$\alpha = 0.265$ $\beta = 0.098$ $\gamma = 1$	0.217	0.527	0.000
3	0.198	0.793	$\gamma = 2.4$	0.161	0.791	0.037	$\alpha = 0$ $\beta = 0$	0.273	0.909	0.138	$\alpha = 0.394$ $\beta = 0.115$	0.165	0.791	0.033	$\alpha = 0.054$ $\beta = 0.042$ $\gamma = 2$	0.161	0.793	0.037	$\alpha = 0.449$ $\beta = 0.449$ $\gamma = 0.1$	0.189	0.789	0.010
4	0.286	0.736	$\gamma = 2.4$	0.258	0.736	0.028	$\alpha = 0$ $\beta = 0$	0.364	0.818	0.113	$\alpha = 1$ $\beta = 0.026$	0.290	0.735	0.004	$\alpha = 0$ $\beta = 0.106$ $\gamma = 2$	0.265	0.740	0.021	$\alpha = 0.72$ $\beta = 0.057$ $\gamma = 1.1$	0.286	0.736	0.000
5	0.327	0.664	$\gamma = 2.5$	0.291	0.657	0.037	$\alpha = 0$ $\beta = 0$	0.364	0.727	0.073	$\alpha = 0$ $\beta = 0.315$	0.327	0.664	0.000	$\alpha = 0$ $\beta = 0.12$ $\gamma = 2$	0.299	0.665	0.028	$\alpha = 0.889$ $\beta = 0.153$ $\gamma = 0.8$	0.327	0.664	0.000
6	0.445	0.596	$\gamma = 1.8$	0.401	0.590	0.044	$\alpha = 0$ $\beta = 0$	0.455	0.636	0.041	$\alpha = 0$ $\beta = 0.257$	0.428	0.600	0.018	$\alpha = 0$ $\beta = 0.029$ $\gamma = 2$	0.401	0.606	0.045	$\alpha = 0.133$ $\beta = 0.35$ $\gamma = 0.6$	0.445	0.596	0.000
7	0.141	0.564	$\gamma = 1.4$	0.097	0.564	0.044	$\alpha = 0$ $\beta = 0.351$	0.186	0.888	0.327	$\alpha = 0$ $\beta = 0.328$	0.139	0.565	0.002	$\alpha = 0$ $\beta = 0$ $\gamma = 2$	0.104	0.634	0.079	$\alpha = 0.078$ $\beta = 0.366$ $\gamma = 0.6$	0.141	0.564	0.000
8	0.25	0.586	$\gamma = 2.2$	0.262	0.579	0.014	$\alpha = 0$ $\beta = 0.249$	0.258	0.722	0.136	$\alpha = 0.473$ $\beta = 0$	0.257	0.585	0.007	$\alpha = 0.051$ $\beta = 0$ $\gamma = 2$	0.260	0.585	0.010	$\alpha = 0.279$ $\beta = 0$ $\gamma = 1.3$	0.255	0.585	0.005
9	0.254	0.827	$\gamma = 3.5$	0.204	0.837	0.051	$\alpha = 0$ $\beta = 0.001$	0.273	0.909	0.084	$\alpha = 0.26$ $\beta = 0.363$	0.254	0.827	0.000	$\alpha = 0.061$ $\beta = 0.28$ $\gamma = 2$	0.254	0.827	0.000	$\alpha = 0.061$ $\beta = 0.28$ $\gamma = 2$	0.254	0.827	0.000
10	0.366	0.699	$\gamma = 2$	0.304	0.724	0.067	$\alpha = 0$ $\beta = 0.065$	0.389	0.813	0.116	$\alpha = 0.091$ $\beta = 0.318$	0.366	0.699	0.000	$\alpha = 0$ $\beta = 0.147$ $\gamma = 2$	0.337	0.745	0.055	$\alpha = 0.946$ $\beta = 0.266$ $\gamma = 0.5$	0.366	0.699	0.000
11	0.331	0.652	$\gamma = 2$	0.354	0.646	0.024	$\alpha = 0$ $\beta = 0$	0.364	0.727	0.082	$\alpha = 0.46$ $\beta = 0$	0.345	0.649	0.014	$\alpha = 0.01$ $\beta = 0$ $\gamma = 2$	0.354	0.650	0.023	$\alpha = 0.374$ $\beta = 0$ $\gamma = 1.1$	0.345	0.649	0.014
12	0.439	0.604	$\gamma = 1.7$	0.451	0.590	0.019	$\alpha = 0$ $\beta = 0$	0.455	0.636	0.036	$\alpha = 0.394$ $\beta = 0.014$	0.439	0.604	0.000	$\alpha = 0$ $\beta = 0$ $\gamma = 2$	0.466	0.604	0.027	$\alpha = 0.889$ $\beta = 0.002$ $\gamma = 0.7$	0.439	0.604	0.000

Table 3.6: Results of the equilibria computation for the original 12 games using the modified concepts

3.6.2 Methodological Formalizations

This section documents the theoretical definitions of cyclic games, the stationary equilibria concepts, the inequity aversion transformations, and the Q measures.

Cyclic Games Taxonomy

We performed an experiment using an inspection game design, following the structure outlined in Selten and Chmura (2008) for cyclic games, illustrated in the matrix in table 3.7. The analyzed 2×2 game, represented in the following standard form, has precisely one completely mixed Nash equilibrium.

		Pl 2	
		L	**R**
Pl 1	**U**	$(\mathrm{a_L} + \mathrm{c_L}, \mathrm{b_U})$	$(\mathrm{a_R}, \mathrm{b_U} + \mathrm{d_U})$
	D	$(\mathrm{a_L}, \mathrm{b_D} + \mathrm{d_D})$	$(\mathrm{a_R} + \mathrm{c_R}, \mathrm{b_D})$

$a_L, a_R, b_U, b_D \geq 0$; $c_L, c_R, d_U, d_D > 0$

Table 3.7: Taxonomy of experimental 2×2 cyclic games

Nash Equilibrium (NE)

The NE formula for computing the equilibrium of cyclic games is given as:

$$p_U = \frac{d_D}{d_U + d_D}, \quad q_L = \frac{c_R}{c_L + c_R}. \tag{3.1}$$

Quantal Response Equilibrium (QRE)

For computing the QRE, let $E_U(q)$ and $E_D(q)$ be player one's expected payoffs for U and D, against a strategy q of player two. Conversely, $E_L(p)$ and $E_R(p)$ are the expected payoffs for player two's L and R against a strategy p for player one. In the QRE, p_U and q_L are defined as:

$$\begin{gathered} p_U = \frac{e^{\lambda E_U(q)}}{e^{\lambda E_U(q)} + e^{\lambda E_D(q)}}, \quad q_L = \frac{e^{\lambda E_L(p)}}{e^{\lambda E_L(p)} + e^{\lambda E_R(p)}}. \\ E'_U(q) = E_U(q) + q_R r, \quad E'_D(q) = E_D(q) + q_R r \end{gathered} \tag{3.2}$$

In this equations system, the mixed-strategy profiles are given as a function of λ, which minimizes the distance of predictions from the observed relative choice probabilities. By incorporating the possibility of errors in human judgment, the QRE is given as a modified version of the NE, introducing noise into the local optimization process.

Action Sampling Equilibrium (ASE)

In the stationary state defined by the outlined game structure, players take a sample of n past choices of the other players and optimize the decision against this sample. To determine the probabilities of the strategies in the action sample equilibrium for sample size n, one determines for each number $k \leq n$, where k is the number of L played and $n-k$ the numbers of R played in the sample of length n, an indicator $\alpha_U(k)$ for the payoff difference $\Delta_1 = (n-k)c_R - kc_L$

$$\alpha_U(k) = \begin{cases} 1 \text{ for } 0 < \Delta_1 \\ \frac{1}{2} \text{ for } 0 = \Delta_1 \\ 0 \text{ else} \end{cases}. \tag{3.3}$$

In the same way for player two for each number $m \leq n$, where m is the number of U played and $n-k$ the numbers of D played in the sample of length n, an indicator $\alpha_L(k)$ for the payoff difference $\Delta_2 = (n-m)d_U - md_D$:

$$\alpha_L(m) = \begin{cases} 1 \text{ for } 0 < \Delta_2 \\ \frac{1}{2} \text{ for } 0 = \Delta_2 \\ 0 \text{ else} \end{cases}. \tag{3.4}$$

The choice probabilities are given for the action sample with a sample size n by:

$$p_U = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) q_L^k \left(1-q_L\right)^{n-k} \alpha_U(k), \quad q_L = \sum_{m=0}^{n} \left(\begin{array}{c} n \\ m \end{array} \right) \left(1-p_U\right)^m p_U^{n-m} \alpha_L(m). \tag{3.5}$$

In the analysis presented here, we tested all ASE sample sizes ranging from 1 to 12.

Payoff Sampling Equilibrium (PSE)

For the PSE, each player takes two samples of equal size for each pure strategy. The sum of payoffs is then compared, and the strategy with the higher payoff sum is selected (if both are equal, then each strategy is selected with the probability of $\frac{1}{2}$), (Osborne & Rubinstein, 1998). Selten and Chmura (2008) introduced The PSE concept as a mixed-strategy combination reflecting this situation. The equations for the payoff sampling with sample length n are derived as follows:

Let k_U be the number of U and k_D the number of D in the sample of player 1 and further

$$\begin{gathered} H_U = k_U \left(a_L + c_L\right) + \left(n - k_U\right) a_R, \quad H_D = k_D a_L + \left(n - k_D\right)\left(a_R + c_R\right) \\ \Phi_1 = H_U - H_D, \\ H_L = m_U b_U + \left(n - m_L\right)\left(b_D + d_D\right), \quad H_D = m_R \left(b_U + d_U\right) + \left(n - m_R\right) b_D \\ \Phi_2 = H_L - H_R. \end{gathered}. \tag{3.6}$$

With the indicator

$$\beta\left(k_U, k_D\right)=\begin{cases}1 \text{ for } 0<\Phi_1\\ \dfrac{1}{2} \text{ for } 0=\Phi_1,\\ 0 \text{ else}\end{cases} \qquad \gamma\left(m_L, m_R\right)=\begin{cases}1 \text{ for } 0<\Phi_2\\ \dfrac{1}{2} \text{ for } 0=\Phi_2\,.\\ 0 \text{ else}\end{cases} \tag{3.7}$$

The system of equations to compute the equilibrium probabilities is defined as:

$$p_U=\sum_{k_U=0}^{n}\sum_{k_D=0}^{n}\binom{n}{k_U}\binom{n}{k_D} q_L^{k_U+k_D}\left(1-q_L\right)^{2n-k_U-k_D}\beta\left(k_U, k_D\right), \tag{3.8}$$

$$q_L=\sum_{m_L=0}^{n}\sum_{m_R=0}^{n}\binom{n}{m_L}\binom{n}{m_R}\left(1-p_U\right)^{m_L+m_R} p_U^{2n-m_L-m_R}\gamma\left(m_L, m_R\right). \tag{3.9}$$

In accordance with the ASE sample size choices, we also analyzed the values between 1 and 12 for PSE. In both ASE and PSE models, increasing the sample size range increases the calculation complexity by increasing the degree of the polynomials.

Impulse Balance Equilibrium (IBE and IBEγ)

In the IBE, the probability of choosing a strategy is taken as a parameter to be adjusted. The authors assumed that the pure strategy maximin is the reference for determining a surplus or a loss. The original game matrix must be transformed to specify the γ-Impulse equilibrium. All payoffs below the surplus s_i

$$s_1=\max\left[\min\left(a_L+c_L, a_R\right), \min\left(a_L, a_R+c_R\right)\right], \tag{3.10}$$

and

$$s_2=\max\left[\min\left(b_U, b_D+d_D\right), \min\left(b_U+d_U, b_D\right)\right] \tag{3.11}$$

Remain unchanged. The surplus of payoffs higher than s_i is reduced by the factor γ between a given and an alternative strategy against a fixed opponent strategy.

The original IBE concept uses $\gamma=2$, based on the loss aversion definition in Kahneman and Tversky (1979). We applied the original method and a dynamic γ-modified version as an optimization model that minimizes the distance to observed probabilities, analogous to the QRE parameter λ. The transformation involves iterating over player i's payoffs x_i, comparing them to security levels s_i, and reducing the surplus over s_i by the factor γ. The transformed payoffs t_i for player p_i are given by:

$$t_i=\begin{cases}s_i+\frac{x_i-s_i}{\gamma}, & \text{if } x_i \geq s_i\\ x_i, & \text{else}\end{cases} \tag{3.12}$$

The transformed game matrix accounts for the combination of ordinary and loss impulses from one pure strategy to another. In the transformed game, the payoff differences c_L, c_R, d_U and d_D are expressed by c^*_R, c^*_L, d^*_L and d^*_D. When a strategy is selected by player i, and it yields a payoff inferior to another option, player i receives an impulse toward the other strategy. The foregone payoff gives the size of the impulse from the unselected strategy in the transformed game. With $x^*_y = x_y\gamma$ for $x_y \in \{c^*_R, c^*_L, d^*_L, d^*_D\}$ the γ-impulse matrix is then given by table 3.8.

		Pl 2	
		L	**R**
Pl 1	U	$(0, d^*_L)$	$(c^*_R, 0)$
	D	$(c^*_L, 0)$	$(0, d^*_D)$

Table 3.8: Impulse matrix

The IBE assumes that player 1's expected impulse from strategy U to D is the same as the impulse from D to U. This logic also applies to payer 2's impulse from L to R and R to L, yielding the following impulse balance equations:

$$p_U q_R c^*_R = p_D q_L c^*_R, \qquad p_U q_L d^*_U = p_D q_R d^*_D. \tag{3.13}$$

The terms on the left and right-hand sides of the first impulse balance equation denote player 1's expected impulse from U to D and player 1's expected impulse from D to U, respectively. If the left-hand side is higher than the right-hand side, player 1 has a stronger impulse from R to D, which consequently decreases q_R and increases q_L, thus creating a tendency in the direction of the impulse balance, as argued by Selten and Chmura (2008). The impulse balance equations yield the following equations for the equilibrium state functions:

$$p_U = \frac{q_L c^*_L}{q_L c^*_L + (1-q_L)c^*_R}, \qquad q_L = \frac{(1-p_u)d^*_D}{p_U d^*_U + (1-p_U)d^*_D}. \tag{3.14}$$

The γ-impulse equilibrium can then be computed by determining the coordinates of the intersection of (p_U, q_L), expressed as:

$$p_U = \frac{\sqrt{c}}{\sqrt{c}+\sqrt{d}} \qquad q_L = \frac{1}{1+\sqrt{cd}}, \tag{3.15}$$

with

$$c = \frac{c^*_L}{c^*_R} \qquad d = \frac{d^*_U}{d^*_D}. \tag{3.16}$$

We shall use IBEγ as an abbreviation for the γ-impulse to distinguish these concepts. We applied both methods with fixed and dynamic parameters analogously.

Inequity Aversion

For a player i with payoffs x, the original model described in Fehr and Schmidt (1999) is defined as:

$$\mathrm{U_i}\left(\mathrm{x_i}, \mathrm{x_j}\right) = \mathrm{x_i} - \max\left[\alpha_i\left(\mathrm{x_i} - \mathrm{x_j}\right), 0\right] - \max\left[\beta_i\left(\mathrm{x_i} - \mathrm{x_i}\right), 0\right], \alpha_i, \beta_\mathrm{i} \in [0,1], \tag{3.17}$$

where *alpha* captures the distaste of player i for disadvantageous inequality in the first nonstandard term, also denominated as "envy," whereas *beta* captures the distaste of player i for advantageous inequality in the final term, or "guilt." This formula is applied to modify the original game matrix prior to equilibria computation.

Euclidean Distance - Goodness of Fit for Model Predictions

For each stationary concept, we calculated the Euclidean distances between the predicted probabilities and the empirical observations to evaluate their predictive success. f_{iUj} and f_{iLj} are the relative frequencies for U and L in the j-th independent subject group. The euclidean distance to the predicted probabilities (p_U, q_L) is calculated as:

$$Q_{ij} = \sqrt{(f_{iUj} - p_U)^2 + (f_{iLj} - q_L)^2} \tag{3.18}$$

Subsequently, the aggregated mean squared distance for a given set of players is defined as follows.

$$Q_i = \frac{1}{s_i}\sum_{j=1}^{s} Q_{ij} \tag{3.19}$$

We used this method to measure the Q values for all aggregations of the data observed in the experiment, considering the overall treatment level, round-based, and individual game-play level.

3.6.3 Technical Remarks

The experiment was programmed using oTree (D. L. Chen, Schonger, & Wickens, 2016). The statistical tests were employed using the statsmodels library (Seabold & Perktold, 2010). Both the experiment and the analysis were executed using Python language. For the equilibria computations, we used Scipy (Virtanen et al., 2020) and Gambit (McKelvey, McLennan, & Turocy, 2023).

Regarding the stationary concept calculations in the context of the introduced models, creating equilibria predictions for a game is primarily a mathematical optimization problem. We aim to comprehend how parameters and other factors produce and affect these predictions. To accomplish this, we organized the calculations of equilibria predictions as grid search problems. In this methodological approach, we define boundaries and increments for all possible combinations of parameters, thus generating a search space. The grid search method iterates through all possible combinations and finds optimal values that satisfy the objective function (LaValle, Branicky, & Lindemann, 2004; Liashchynskyi & Liashchynskyi, 2019), in our case defined by our measure of predictive success Q.

The modified models had three possible parameters: γ, α, and β. We defined the search space by defining boundaries and increment steps for each. The boundaries for α and β were set to 0 and 1, with increment steps of 0.001, whereas for γ, we had it between 0.1 and 12 with increments of 0.1. It is important to outline here that searching through large grids can be a complex computational task. If we think of the model with 3 parameters, $IBE_\gamma \circ IA_{\alpha,\beta}$, we had a search space of $1001*1001*120$, which searches through 120240120 pairs of predicted probabilities for (p_U, q_L) for a given game. Our implementation was done using Python programming language. The code was written for multi-thread tasks to handle large search spaces using open-source just-in-time compilers from the Numba library (Lam, Pitrou, & Seibert, 2015). During the grid search process, IBE-derived methods may generate games that don't follow cyclic game rules (table 3.7), making it impossible to calculate impulse equilibrium. To address this, we check if games remain cyclic after applying inequity aversion and impulse transformations. If not, the algorithm skips predictions and moves to the next case.

3.6.4 Experiment Design Details

In this section, we added the actual experiment designs, which the subjects interacted with, containing examples of framed and unframed variants.

The experiments were conducted in Kleve, Germany's Rhine-Waal University of Applied Sciences experimental laboratory.

Introduction - How matrix games work

We'll give a short introduction on how the games -called Matrix games- are played: Matrix games describe two person (Player 1 & Player 2) decision situations.

Example:

		Player 2	
		A	B
Player 1	A	(2, 2)	(0, 3)
	B	(3, 0)	(1, 1)

In this example, the possible strategies for each player are named A & B and each player can pick one of these two strategies

The corresponding payoffs are given in the brackets. The first number is the payoff for player 1 and the second number is the payoff for player 2. So for example, (5, 2) denotes a payoff of 5 for player 1 and 2 for player 2.

Therefore, the final payoff depends not only on the own individual decision, it is also based on the decision of the other player. But player 1 and 2 choose their strategies without knowing the decision of the other player.

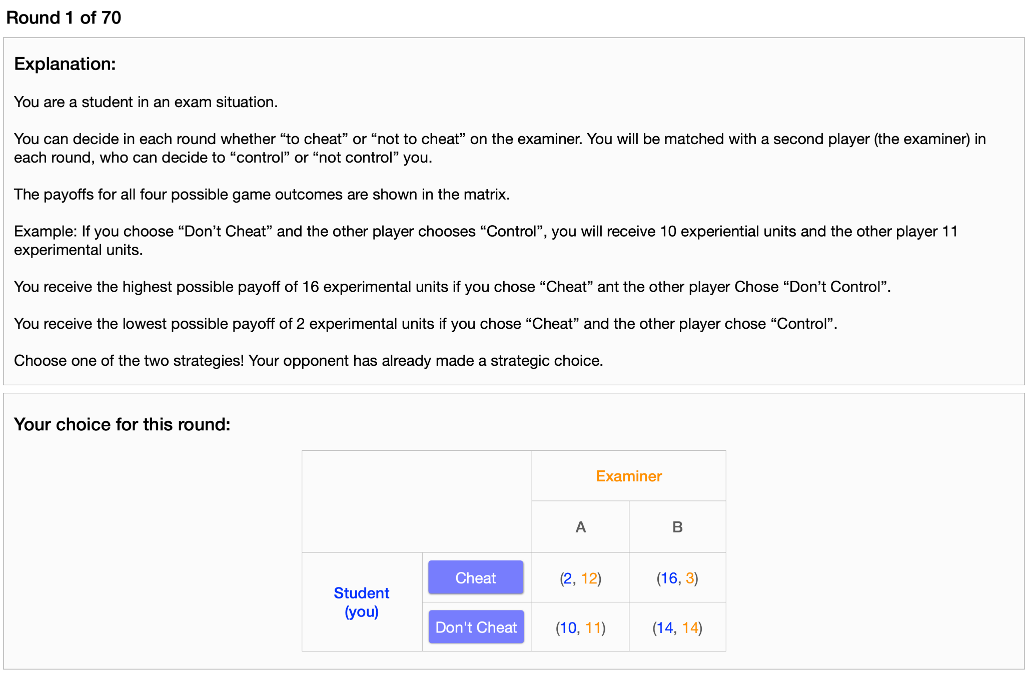
Example: If player 1 chooses strategy A, and player 2 strategy B, we find this situation:

		Player 2	
		A	B
Player 1	A	(2, 2)	(0, 3)
	B	(3, 0)	(1, 1)

The payoff for player 1 is 0 and for player 2 it is 3.

You will play against another human player, who will be chosen randomly in the beginning, but will remain the same for all subsequent rounds.

Figure 3.14: Introductory explanation



Round 1 of 70

Explanation:

You are a student in an exam situation.

You can decide in each round whether "to cheat" or "not to cheat" on the examiner. You will be matched with a second player (the examiner) in each round, who can decide to "control" or "not control" you.

The payoffs for all four possible game outcomes are shown in the matrix.

Example: If you choose "Don't Cheat" and the other player chooses "Control", you will receive 10 experiential units and the other player 11 experimental units.

You receive the highest possible payoff of 16 experimental units if you chose "Cheat" ant the other player Chose "Don't Control".

You receive the lowest possible payoff of 2 experimental units if you chose "Cheat" and the other player chose "Control".

Choose one of the two strategies! Your opponent has already made a strategic choice.

Your choice for this round:

		Examiner	
		A	B
Student (you)	Cheat	(2, 12)	(16, 3)
	Don't Cheat	(10, 11)	(14, 14)

Figure 3.15: Framed design example, applied to the student player

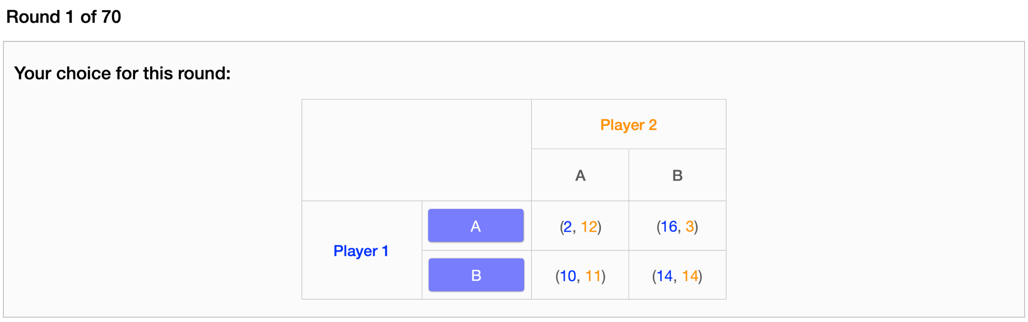


Figure 3.16: Unframed design example, applied to the student player

4. Understanding Dark Personality Traits and Strategic Choices in an Inspection Game: Insights from Machine Learning and Causal Inference

Authors

Vinícius Ferraz, Leon Houf, Thomas Pitz & Christiane Schwieren

Abstract

This paper analyzes the dynamics of human behavior in strategic interactions, focusing on the influence of dark personality traits within the framework of an inspector game. Utilizing a rigorous experimental design, participants engaged in one-shot and repeated game versions under two distinct framing conditions: neutral and information-loaded. The analysis employed a combination of traditional statistical analyses complemented by advanced machine learning techniques, including Random Forest and Causal Forest models. Key findings highlight the prevailing role of game mechanics in decision-making, with the D-factor surfacing as a significant secondary influencer. The inspector profile shows higher retaliation trends and consistently exhibited pronounced treatment effects, especially under framed conditions. The results offer a thorough understanding of the interplay between individual attributes, game dynamics, and strategic choices, contributing valuable insights into behavioral economics and strategic decision-making.

Keywords

Behavioral Game Theory, Dark Personality, Machine Learning, Causal Inference, Inspection Games

4.1 Introduction

In strategic decision-making, a comprehensive understanding of the factors influencing human behavior is essential C. Camerer (2011). Traditional economic theories focus on rationality, (Kahneman & Tversky, 1979, 1984) and highlight the impact of cognitive biases and emotions. Additionally, research on cooperation Axelrod and Hamilton (1981), fairness Fehr and Gächter (2002); Fehr and Schmidt (1999), and social preferences Charness and Rabin (2002) underscores the diverse influences in strategic interactions. Building on this perspective, this paper analyzes the influence of dark personality traits within the framework of an inspector game. By employing a combination of traditional statistical analyses and advanced machine learning techniques, our study explores the relationship between game mechanics, contextual framing, and individual traits like the Dark Factor of

Personality (D-factor), which measures characteristics associated with manipulative, self-centered, and malevolent behaviors. Our research aims to provide a deeper understanding of how these elements converge to influence strategic behavior in decision-making scenarios laden with asymmetric information and non-rational elements.

Inspection games model a scenario where an inspector aims to detect non-compliant behaviors of an inspectee, often within resource constraints. In this context, cheating refers to deceiving or manipulating to gain an advantage despite the risk of detection. While this game provides a robust framework for studying behaviors like dishonesty, retaliation, and punishment, the influence of dark personality traits remains largely unexplored. When combined with framing and dark traits, the game's particular dynamics introduce new directions in behavioral research. Furthermore, our methodological approach, which integrates traditional statistical analyses with advanced machine learning, offers comprehensive insights into behavioral patterns and the causal relationships between strategic choices, context, and individual traits.

In the experiment's design, we incorporated both one-shot and repeated variations of an Inspection Game, enabling in-depth examinations of strategic behavior statically and over time. Participants were randomly assigned to one of two framing schemes: neutral or information-loaded, introducing contextual elements to the decision-making process. Post-game questionnaires assess individual characteristics like dark personality traits and propensity for dishonesty, serving as pivotal variables in our analysis. This comprehensive design allows us to draw causal inferences and provides a robust framework for examining interactions between personality, context, and decision-making. Our analytical approach focuses on participants' strategy selections, specifically their tendencies to "cheat" or "control," and extends to correlations and regressions. Subsequently, it expands to incorporate Random Forest and Causal Forest models, facilitating a deeper understanding of behavioral patterns and their underlying causal relationships.

Our findings underscore the dominant influence of game mechanics on decision-making, with the D-factor emerging as a secondary determinant. The inspector profile consistently exhibited more pronounced treatment effects and higher retaliation tendencies. At the same time, the causal forest model highlighted the nuanced impacts of the dark factor, age, and gender on treatment outcomes. The layered nature of our findings reveals the interwoven relationships between rational decision-making, behavioral context, and individual characteristics.

In terms of practical contributions, we aim to advance the understanding of human behavior and strategic interactions through a methodologically robust and diverse framework. The insights gleaned from our findings offer a meticulous comprehension of cheating and controlling behavior, which has implications for policy formulation and business ethics. Among the practical applications of our research are the potential for detecting cheating behaviors and designing targeted interventions based on individual risk factors.

4.2 Conceptual Background

This section examines two core elements of the study: Inspection Games, dark personality traits, and their role in strategic interactions and cheating. We discuss the use of Inspection Games in game theory, explore the Dark Factor of Personality (D) and its impact on decision-making, and highlight empirical evidence linking dark traits to strategic and cheating behavior.

4.2.1 Inspection Games

This paper employs an experiment using a variation of the Inspector Game, which belongs to the Cyclic class. Cyclic games, as defined in Selten and Wooders (2001), are a type of game theory model designed to account for recurring situations where the same scenario may be repeated indefinitely, potentially with different players. They are characterized by having payoff matrices where the best replies of the players cycle, resulting in a situation where there is no pure-strategy Nash equilibrium. Instead, these games often have mixed-strategy equilibria, where players utilize all their possible strategies with a given probability (Kaplansky, 1945), and the game dynamics tend to cycle around them.

The game involves a sequential play between an inspector and an inspectee. The inspector aims to detect illicit activities by the inspectee through inspecting actions. However, due to limited resources, continuous inspection is unfeasible. This formulation seeks an equilibrium representing the optimal strategies for both players, considering their resources and the other's strategy. Often used in international relations modeling, such as arms treaty enforcement, inspector games encapsulate strategic interactions under asymmetric information. Key contributions include works by Avenhaus et al. (1996, 2002). Early definitions are attributed to studies by Dresher (1962) and Maschler (1966). Inspection games have found application in various fields including law enforcement (Andreozzi, 2004; Ferguson & Melolidakis, 1998; Rauhut, 2009), environmental policy (Holler, 1993), fraud detection (Kolokoltsov, 2010), corporate bonuses and fines (Nosenzo et al., 2010), public transport controls, arms control (Avenhaus, 2004), to name a few.

The Inspector game can provide a framework that captures the essence of a moral decision-making scenario, where players must weigh personal gain against their choices' social and moral implications. This game-theoretic approach is especially suitable for examining dark personality traits, as it mirrors real-world contexts where individuals possessing these traits may take advantage of others for their benefit.

4.2.2 Dark Personality Traits

Dark personality traits encompass behaviors that are harmful and self-centered. These individuals often prioritize their own needs, showing little empathy or warmth towards others. They can be deceptive, misleading others for their own benefit, and may intentionally harm others to achieve their goals (Rogoza, Kowalski, Saklofske, & Schermer, 2022).

While the dark triad of personality, comprising Machiavellianism, Narcissism, and Psychopathy, is one of the most studied constructs in this field (Paulhus & Williams, 2002), dark traits extend beyond these three. Recognizing the need for a comprehensive framework, Moshagen, Hilbig, and Zettler (2018) introduced the Dark Factor of Personality (D-factor). This value combines the core dispositional tendencies underlying various aversive personality traits. D is characterized by an individual's inclination to maximize their utility, often at the expense of others, and is supported by beliefs that justify such behaviors. It consolidates various dark traits, including but not limited to Machiavellianism, Narcissism, Psychopathy, Amorality, Egoism, Greed, Sadism, and Spitefulness, into a single measure, providing a holistic perspective.

The decision to employ the D-factor stems from its comprehensive coverage of dark traits and the methodological advantage of consolidating them into a singular score. The association of dark personality traits with behaviors such as callousness, dishonesty, selfishness, and risk-taking is well-established (Crysel, Crosier, & Webster, 2013; Jones & Neria, 2015; Moshagen et al., 2018). Given the relevance of these traits to our study, we employ the D-factor measurement questionnaire as proposed by Moshagen, Zettler, and Hilbig (2020).

4.2.3 Dark Personality Traits, Strategic Interaction, and Cheating - Related Work

The literature reveals a consistent relationship between dark personality traits and various forms of dishonest or strategic behavior. Studies by K. Smith, Emerson, Haight, and Wood (2022), G. J. Curtis, Correia, and Davis (2022), and Esteves, Oliveira, de Andrade, and Menezes (2021) collectively indicate that these traits are significant predictors of academic misconduct among students, with some variations based on cultural and academic contexts (Cheung & Egan, 2021; J. Zhang, Paulhus, & Ziegler, 2019).

In strategic games like the Prisoner's Dilemma, dark traits link to more selfish or strategic behaviors. Deutchman and Sullivan (2018) and Lainidi, Karakasidou, and Montgomery (2022) found that individuals with high Dark Triad scores were more likely to betray their partners, especially in loss-framed and non-social contexts. Lopez, Calvo, and Torre (2022) extended this to show that Dark Triad traits influence decision-making but may not affect susceptibility to anchoring bias.

In competitive settings, S. R. Curtis et al. (2021) found that high Machiavellianism led to better performance, whereas high levels of narcissism and psychopathy resulted in poorer outcomes. This result aligns with Timmermans, Caluwé, and Alexopoulos (2018)'s findings that dark traits like psychopathy correlate with immoral behaviors such as infidelity on online dating platforms.

This body of work underscores dark personality traits' complex but consistent influence on dishonest behaviors and strategic decision-making across various contexts.

4.3 Experiment Design

We use a between-subject design, randomly assigning participants to one of two treatments with distinct framing schemes. This approach enables causal behavior comparisons across treatments (Charness, Gneezy, & Kuhn, 2012). At the beginning of the experiment, participants are also randomly paired. In the repeated game, these pairs remain consistent from start to finish. However, in the one-shot game, participants submit strategies individually. Post-game, participants completed the dark factor survey. Additionally, the one-shot group answered five dishonesty-related questions, enhancing and diversifying findings due to the game's lack of a time dimension. Further details on the game matrix, framing schemes, and questionnaires will follow in subsequent sections.

4.3.1 Game Matrix

The game consists of a 2×2 strategic-form version of the inspection game, designed following Selten and Chmura (2008)'s taxonomy for cyclic games. The game models an exam situation, introducing the strategic interaction between an inspector and a student. The Student has the strategy options "cheat" or "not cheat," whereas the inspector can choose between "control" and "not control." Figure 4.1 displays the game matrix representing this scenario, with player 1's (row) payoffs representing the inspector profile and player 2's (column) the student.

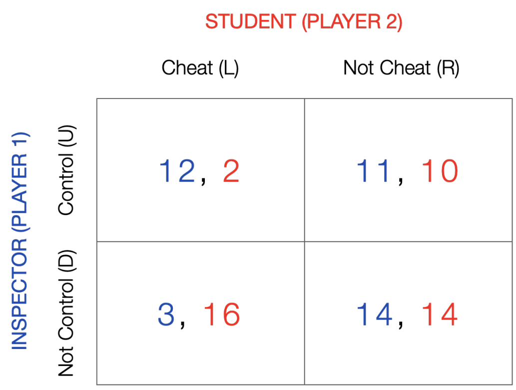


Figure 4.1: Base Inspector Game Matrix

Considering real-life scenarios, we model "cheating" as an easier, atypical behavior, contrasting it with a "control" condition that entails effort or cost. The game's payoffs embody this, leaning towards the $(\neg control, \neg cheat)$ outcome (where "$\neg$" denotes negation-based strategies). This model suggests cheating is often simpler than its prevention or control. Hence, the game's asymmetric payoffs express the student player's preference hierarchy:

$$(\neg Control, Cheat) \succ (\neg Control, \neg Cheat) \succ (Control, \neg Cheat) \succ (Control, Cheat), \tag{4.1}$$

Suggesting that cheating, when not controlled, is the most preferred outcome. Complementarily, we designed the following preference relations for the inspector player:

$$(\neg Control, \neg Cheat) \succ (Control, Cheat) \succ (Control, \neg Cheat) \succ (\neg Control, Cheat), \tag{4.2}$$

where the preferred outcome is no cheating and no control. The preference relations are reflected in the players' best response correspondences, introducing a unique, completely mixed Nash equilibrium. As such, our model's equilibrium predicts that the inspector will opt for control with a probability $p_U = 0.2$ and choose not to control with a higher probability $p_D = 0.8$. On the other hand, the Student is anticipated to cheat with probability $q_L = 0.25$ and decide against cheating with probability $q_R = 0.75$. This equilibrium reflects the inherent tendencies and strategic choices of both the inspector and the Student within the game framework (for details on the computation of the Nash equilibrium in cyclic games, see appendix 4.7.2).

4.3.2 Information-Loaded and Neutral Frames

Research has shown that framing choices can significantly influence decision-making outcomes, even when the underlying information remains the same. This importance of cognitive biases and heuristics challenges the idea of pure rationality being our primary decision driver (Tversky & Kahneman, 1981). How information is presented can alter perceptions and decisions through cognitive and motivational effects (Levin, Schneider, & Gaeth, 1998). Framing effects on individual and group behavior in 2×2 game experiments have been substantiated in studies such as Brewer and Kramer (1986) and Andreoni (1995).

Among different types of frames, the subset of interest in our context comprises neutral and information-loaded frames, as studied in Abbink and Hennig-Schmidt (2006). This approach is designed to control for potential biases or preconceptions related to the task, particularly given that our experiment involves actions with potentially negative ethical implications. We employ neutral and information-loaded frames, referred to as "unframed treatment" and the "framed treatment," respectively. The unframed treatment presents the tasks context-free, allowing participants to select generic strategies. On the other hand, the framed treatment adds context with an explanation of the game and named strategies. Other studies with this framing scheme are presented in Dufwenberg et al. (2011), Abbink et al. (2002), and Abbink and Hennig-Schmidt (2006). For visualizing how participants saw the different versions of the game, see appendix 4.7.4.

4.3.3 Dark Traits and Dishonesty Measurements

In the last stage of the experiment, we employ the 35-questions version (D35) of the D-factor questionnaire (Moshagen et al., 2020) to assess the underlying personality traits that might influence participants' behavior in our game scenario. The quantification of dark traits can help understand how this behavioral predisposition interacts with the context and participants' decisions, particularly regarding dishonest, selfish, or punitive behavior. A complete overview of the questions is in table 4.14, appendix 4.7.4.

Participants of the one-shot game answered five additional questions to assess their propensity for dishonest behavior. Following the randomized response scheme introduced by Warner (1965) and later modified by Greenberg, Abul-Ela, Simmons, and Horvitz (1969), participants were randomly divided into two groups to ensure anonymity and encourage truthful responses. One group responded directly to questions. The other group was instructed to flip a coin before responding: if the coin showed "heads," they should answer "yes"; if "tails," they should answer truthfully. This coin-flip method aimed to ensure participants felt their responses were anonymous. Further details are provided in appendix 4.7.4.

Based on the outlined design, we posit the following: Dark personality traits, as measured by the D-factor, are hypothesized to influence game behavior, with individuals displaying higher dark traits likely exhibiting more control or cheat strategies. In addition, whether neutral or information-loaded, the game's framing is expected to influence participants' decisions. However, the mechanics of the game (incentive schemes, feedback, repeated interaction, etc.) are anticipated to play an important role in shaping decision-making, regardless of individual traits or framing effects.

4.4 Machine Learning Analysis: Models, Explainability, and Performance Metrics

In recent years, machine learning methods have gained significant traction in tackling complex economic problems. C. F. Camerer (2018) highlights its predictive accuracy and adaptability strengths, especially when dealing with complex human behavior. These features enable more effective personalization and foster cross-disciplinary insights, thereby contributing to a nuanced understanding of human decision-making across various domains. Complementarily, Athey (2018) notes machine learning's robustness in causal inference, particularly in scenarios involving numerous covariates.

To navigate the complex landscape imposed by our experiment, We leverage traditional statistical methods and advanced tools like random forests. We also use causal forests to distinguish the effects of treatment, game mechanics, and personal traits. To interpret the machine learning models, we turn to Shapley values. The subsequent subchapters delve into these methods in detail.

4.4.1 Random Forest

Random Forest is a machine learning algorithm proposed by Breiman (2001), inspired by earlier work in shape recognition with multiple randomized trees by Amit and Geman (1997). It comprises a collection of decision trees, which are simple models that split data into branches at each step. Each tree in the Random Forest is trained on a random subset of the data and makes predictions. The final result of the Random Forest is then made by taking the majority vote (for classification tasks) or the average (for regression tasks) of the trees' predictions. This ensemble approach helps to improve prediction accuracy and robustness. Random Forests excel in managing complex datasets with many variables, adeptly handling interactions between them. This algorithm stands out for its capability to process high-dimensional data and offers built-in error estimation and variable importance metrics (Cutler, Cutler, & Stevens, 2012). A generalized formalization of the Random Forest classifier is documented in appendix 4.7.2.

4.4.2 Causal Forest and Double Machine Learning (DML) Estimators

In this analysis step, we aim to disentangle the treatment effects and the influence of covariates. The goal is to assess how these two sets of factors jointly influenced the participants' behavior. We employ machine learning combined with causal inference techniques to achieve this. Specifically, we focus on estimating the Conditional Average Treatment Effects (CATE), which represent the expected disparity in outcomes between treated and untreated groups, conditioned on a set of observed covariates (Athey & Imbens, 2016). According to Lechner (2023), Causal Forests are particularly effective for their flexible estimation of causal effects at multiple levels of aggregation, allowing for a more nuanced understanding of complex behavioral phenomena.

The concept of causal forests, proposed by Wager and Athey (2018), is an extension of Breiman (2001)'s random forest designed to estimate CATE using causal inference. The causal forest method pairs "honest" trees with the subsampling mechanism of random forests to estimate heterogeneous treatment effects. This algorithm constructs a random forest where each tree is "honest," meaning the data splits into two parts: one for constructing the tree (choosing splits) and the other for estimating treatment effects within the resulting leaves. This separation of data avoids overfitting and allows for valid statistical inference. Meanwhile, the random forest's subsampling helps reduce variance and improve the model's predictive performance. The transition from a Random Forest to a Causal Forest primarily focuses on the tree's target variable and splitting criterion. In Random Forests, the trees aim to minimize classification or regression errors (classification, in our case), but in Causal Forests, they are designed to estimate causal effects. For applications of this method, see Athey and Wager (2019); Davis and Heller (2017).

Complementarily, Double Machine Learning (DML) is a method introduced in Chernozhukov et al. (2017), which integrates decision tree ensembles with the DML approach to robustly estimate localized causal treatment effects in the presence of high-dimensional covariates while accounting for confounders. The name "double" relates to the involvement of two sets of machine learning predictions. The first model (model T) predicts the treatment T given covariates X, yielding a propensity score or probability of receiving the treatment based on observed characteristics. The second model (model Y) predicts the outcome Y given covariates X and cofounders W, highlighting the expected outcome based solely on observed characteristics without considering the treatment. Using the predictions from the two models, the DML model computes the residuals $\tilde{T}$ and $\tilde{Y}$ by subtracting predicted values from the observed treatment and outcome, respectively. These residuals represent the component of treatment and outcome that the observed covariates cannot explain.

The DML version of causal forest integrates generalized random forests (Athey, Tibshirani, & Wager, 2018) with DML estimators, allowing for a nuanced analysis in high-dimensional covariate spaces and heterogeneity in treatment effects. The DML technique employs auxiliary machine learning models to adjust for confounders, ensuring unbiased treatment effect estimation. Given the potential presence of omitted confounders, utilizing a double machine-learning approach is wise. This method combines the robustness of Double Machine Learning with the adaptability of causal forests, enhancing the reliability of treatment effect estimates. For a detailed mathematical breakdown of the Causal Forest DML model, see appendix 4.7.2.

Causal forests with DML estimators have been applied to various fields, such as medicine Mizuguchi and Sawamura (2023), marketing Ellickson, Kar, and Reeder III (2023), ecological monitoring Fink et al. (2022), and finance (Wasserbacher & Spindler, 2022).

4.4.3 Model Evaluation and Performance Metrics

Model evaluation is vital for both prediction and extracting insights. To validate our findings, we examined the performance of the employed two machine learning models: Random Forest Classifier and Causal Forest DML. We partitioned our datasets into 80% for training and 20% for testing. While the Random Forest model's evaluation is direct using the test set, the Causal Forest DML model requires a more intricate approach due to the inherent complexities of causal inference, such as counterfactuals and potential unobserved confounders.

For the Random Forest Classifier, we employed five popular performance metrics: Accuracy, Precision, Recall, F1 Score, and ROC-AUC (Fawcett, 2006; Powers, 2020; Sokolova & Lapalme, 2009). Accuracy calculates the fraction of correctly classified instances, Precision gauges the model's skill in identifying true positives among those it classifies as positive, Recall determines how well the model recognizes positive instances among all real positives, F1 Score balances Precision and Recall, and ROC-AUC assesses the model's discriminative ability across varying thresholds.

In causal machine learning, the causal effect estimation is particularly challenging for model evaluation because it involves unobservable counterfactuals. Therefore, goodness-of-fit metrics like Mean Squared Error (MSE) and R2 are often used, involving only the observed data (Machlanski, Samothrakis, & Clarke, 2023). We utilized two primary evaluation metrics to assess the performance of our Causal Forest DML model: Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). These metrics have been adapted from the field of regression and are particularly suitable for tasks involving the estimation of Average Treatment Effects (ATE), as outlined in Cheng et al. (2022).

4.4.4 Model Selection, Training, and Tuning Strategies

Individualized models were developed for repeated and one-shot game scenarios to understand the multifaceted game dynamics and players' motivations. These models catered to each player type, and random and causal forests were built separately for each dataset.

The double machine learning model has two first-stage models, T (treatment) and Y (outcome), which must be selected and tuned separately. This stage is also crucial for the quality of the final model's results, as demonstrated by Machlanski et al. (2023). Since both our treatment variable and the outcome are binary, this problem naturally falls within the scope of classifiers. Identifying appropriate models for complex datasets involves experimentation and evaluation. We included four model types in the selection process, focused on popular and widely applied models for classification problems: Random Forest (RF), Gradient Boosting Machines (GBM), Gaussian Naive Bayes (NB), and the Multi-layer Perceptron (MLP) classifier. In all cases, we applied cross-validation techniques during model training stages to prevent overfitting while ensuring robustness and reliability. Details on the model descriptions, tuning, and cross-validation are documented in appendix 4.7.3. The systematic first-stage model selection procedure is documented in 4.7.3. Technical choices, parameter values, and details regarding the DML estimation are documented in appendix 4.7.3.

4.4.5 Shapley Values for Machine Learning Explainability

The Shapley Value, originally a cooperative game theory concept by Shapley et al. (1953), denotes a player's average marginal contribution across all player combinations. Adapted to machine learning as Shapley Additive exPlanations (SHAP) by S. M. Lundberg and Lee (2017), it measures feature importance in prediction models. Each feature's SHAP value indicates its average influence on predictions across feature combinations. S. M. Lundberg et al. (2020) extended SHAP to tree-based models, including random and causal forests, offering consistent interpretability for these different methods. For a comprehensive exploration of the SHAP concept, see Sundararajan and Najmi (2020) and appendix 4.7.2.

4.5 Results

This section presents the findings of our experimental analysis. Based on our hypotheses outlining the influence of the game mechanics, the D-factor, and the framing on decisions, we initiate the discussion by examining the sample's characteristics and the behaviors participants exhibited during the game. Subsequently, we incorporate personal variables into our analysis, exploring their relationships through correlation and regression analyses. Finally, we turn to machine learning applications, starting with insights from the Random Forest model and then moving on to the causal inference models.

4.5.1 Samples Information

We conducted experiments using two distinct samples. The first sample involved a repeated game played over 70 periods, while the second sample engaged in a one-shot game. The details about the two samples are summarized in table 4.1. In the repeated game, participants received compensation based on their performance, with each point earned equivalent to 0.02€. On average, participants received a payment of 16.75€ for their participation. In contrast, for the one-shot game, participants had the opportunity to enter a raffle with a chance to win one of three €50 gift cards. It's worth noting that all participants in both experiments were students from Rhine-Waal University of Applied Sciences located in Kleve, Germany.

Attribute	Repeated Game	One-Shot Game
Experiment Type	Laboratory	Online
Participants (N)	100	526
Treatment Split	Framed: 52, Unframed: 48	Framed: 275, Unframed: 251
Gender Distribution	54% Female, 43% Male	47% Male, 48% Female
Nationality	50% German, 50% Diverse	41% German, 59% Diverse
STEM Field Share	32%	53%
Age Distribution		
- 18-24 years	54%	58%
- 25-34 years	40%	37%
- 35-44 years	6%	3%
- Other or non specified	0%	2%
Average Session Length	40 minutes	N/A

Table 4.1: Repeated and One-Shot Samples Comparison

The repeated game experiment was conducted initially. However, we later introduced a one-shot sample to ensure a representative distribution of characteristics, especially for the D-factor. This addressed the lack of extreme values observed in the repeated sample (shown in figure 4.9, appendix 4.7.1). The one-shot approach allowed for an expansive yet cost-effective sample, given the constraints of conducting a large lab experiment

4.5.2 Game Behavior and Statistical Analysis

We comprehensively analyzed game behavior in line with our hypothesis that emphasizes the game's incentives as a pivotal influence on decision-making. This foundational overview serves two purposes: firstly, to provide a clear understanding of the observed game dynamics, and secondly, to set the groundwork for examining its interactions with dark personality traits.

Our analysis began with statistical tests comparing the samples, emphasizing the variable *"s1"* representing the action of choosing the first strategy: "cheat" for students and "control" for inspectors. This variable will be a central point throughout the results section. For the repeated game, we assessed players' aggregate $s1$ choices over 70 periods using Mann-Whitney U (Mann & Whitney, 1947) and permutation (Fisher, 1936) tests, given non-normal distributions. In the one-shot game, we applied permutation, Z-tests, and Fisher's exact test to compare binomial samples. Preliminary results showed no significant differences in strategy by treatment framing or gender. However, there was a statistically significant variation in the inspector's choices in the one-shot game (with the framed sample higher than the unframed). Detailed outcomes of the tests are in table 4.7, appendix 4.7.1.

Upon analyzing the participants' behavior in the repeated game, the observed equilibrium probabilities (p_U, q_L) were identified as $(0.46, 0.13)$ for the framed version and $(0.34, 0.11)$ for the unframed version, taking into account the mean relative frequencies across the game. Meanwhile, the one-shot game's equilibrium probabilities stood at $(0.65, 0.15)$ for the framed version and $(0.47, 0.21)$ for the unframed. In comparing these findings with the

Nash equilibrium predictions of $(0.20, 0.25)$, there were notable deviations, particularly regarding the inspector's strategy selection [1]. Player 1's (the inspector's) choices were consistently higher than the Nash predictions, and Player 2's consistently lower, suggesting a potential predisposition or bias in their decision-making process. Furthermore, the framing of the game appeared to exert influence here. In both game formats, the framed treatment showed a trend of higher $s1$ probabilities for Player 1, indicating a potential context influence. These trends can also be observed by computing the absolute means of selected strategies across treatments and participants, plotted in figure 4.2.

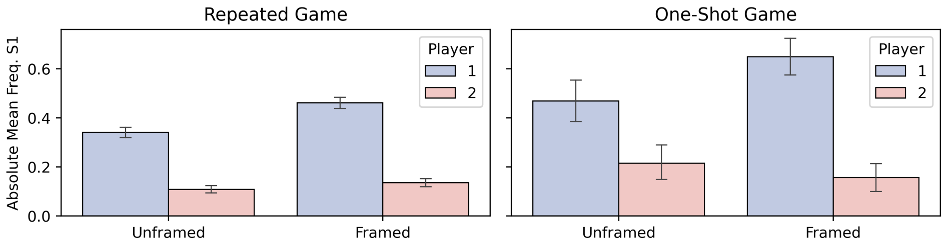


Figure 4.2: Absolute Mean Frequencies of Choosing Strategy 1 (s1).

The analysis of the repeated game behavior is encapsulated in figure 4.3. Starting from the top and moving towards the bottom: First, the time series frequencies of playing strategy 1 reveal that both players initially leaned towards their first strategies, but this preference stabilized after about ten periods, indicating a shift towards coordination with their counterparts. Second, the distribution of payoffs shows comparable mean payoffs for both players, although Player 1 and the unframed treatment in general had slightly higher averages. Third, scatter plots correlating total payoffs with frequencies $s1$ indicate that Player 1 consistently chose $s1$ more often. Third, strong negative correlations between total payoffs and frequent use of $s1$ for both player types outlined that making this choice frequently was not advantageous. Fourth, lower correlation values when examining strategy changes across the game indicate that inconsistency in strategy choices leads to lower payoffs. These findings imply that neither a predisposition to choose $s1$ nor frequent strategy changes optimize payoffs, and the game's framing subtly influenced these behaviors.

The results from the analysis of the one-shot game are available in figure 4.4. The heatmaps (top) reveal distinct patterns in strategy choices between the framed and unframed treatments. In the framed version, the most frequent outcome was Player 1 choosing strategy 1 and Player 2 opting for strategy 2 (54% of the time). However, while this outcome remained prevalent in the unframed treatment, the 2,2 outcome was equivalently frequent (38-40% of the time). Here, we can observe a change not just in individual choices but also in interactions. Regarding payoffs, both treatments show comparable distributions, with slight variations between treatments and players. The framing did not significantly

[1] The scope of this paper will not dive deeper into equilibria computation. McKelvey and Palfrey (1995); Osborne and Rubinstein (1998); Selten and Chmura (2008) discussed the lack of predictive power of the Nash equilibrium concept in cyclic games and proposed alternatives.

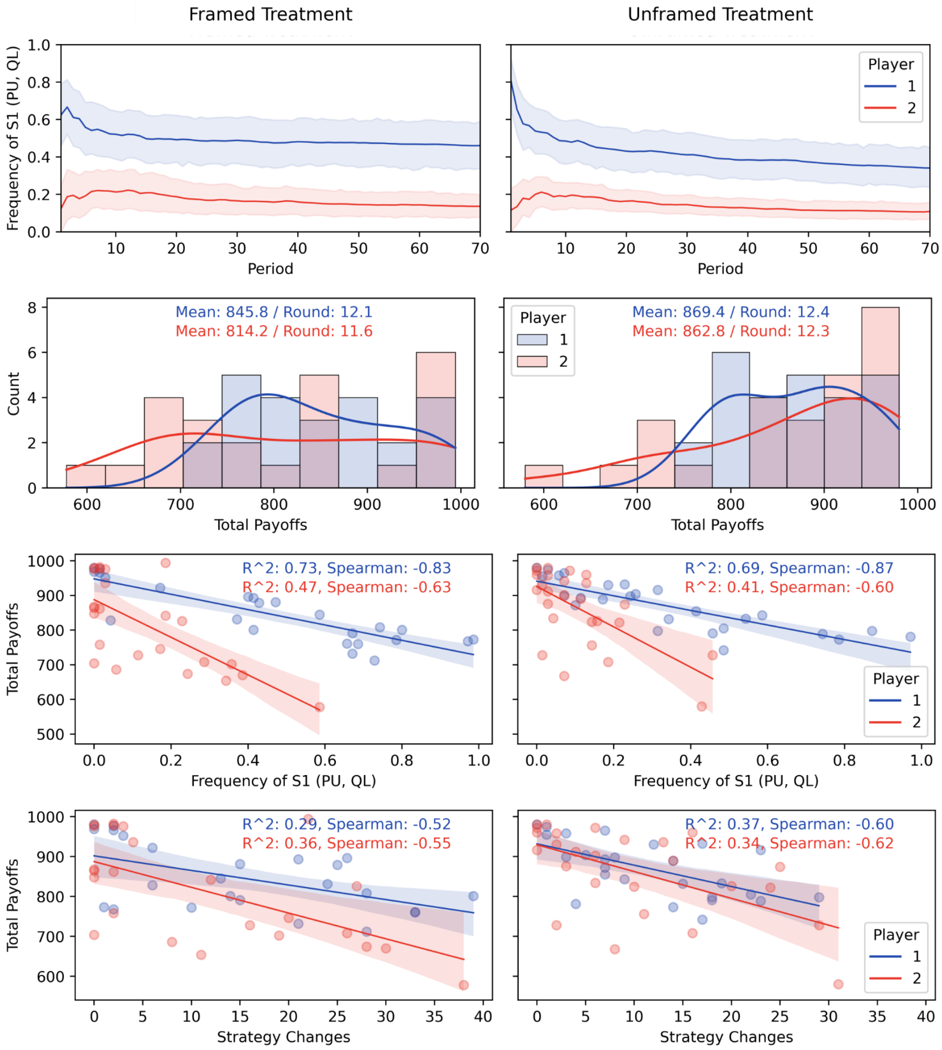


Figure 4.3: Repeated Game Analysis, from top to bottom: (1) time series frequencies of playing $s1$; (2) payoffs distribution; (3) total payoffs vs. frequencies of $s1$ and (4) total payoffs vs. total strategy changes in the game.

impact the overall payoffs but did subtly affect the balance between players.

In analyzing game participant behavior, we observed deviations from traditional game predictions. While strategy variations across treatments were evident, the impact of framing was subtle. The following sections explore further the complexities influencing human strategic behavior.

4.5.3 Dark Factor and Personal Data - Correlations and Regressions

In the following analyses, we consider not only the effects of framing but also the dark personality score D-Factor, dishonesty scores (for the one-shot game), and other personal dimensions such as Age, Gender, and Education Background (STEM). We converted categorical values like Female gender, STEM, and framing into binary indicators. We also divided the data into three categories: the general game context and each player type. Given the distinct motivations and differences between the two player profiles, analyzing

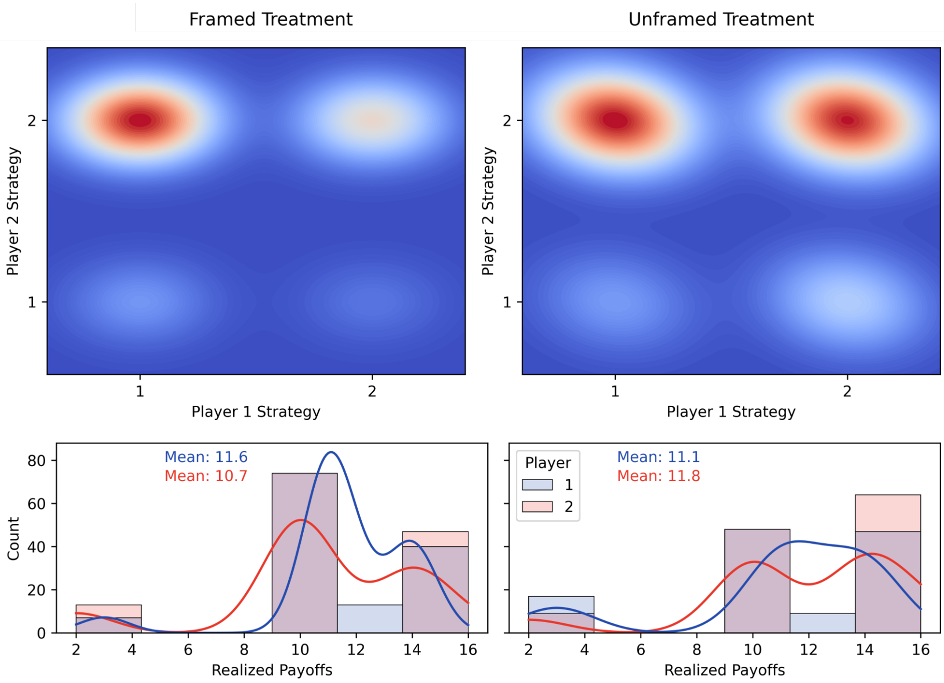


Figure 4.4: One-Shot Game Analysis, from top to bottom: (1) Heatmap of the distributions of the game outcomes and (2) payoff distributions.

them separately and within the broader game dynamics is beneficial to provide insights into behavioral patterns across different scenarios. Figure 4.5 presents the Spearman correlation coefficient matrices in line with this approach. The correlation coefficients between all variables are documented in 4.7.1.

	RG Game	RG P1 (Inspector)	RG P2 (Student)	OS Game	OS P1 (Inspector)	OS P2 (Student)
Age	0.019	0.029	-0.023	-0.018	-0.028	0.010
D-factor	0.140	0.150	0.071	0.081	0.008	0.201
Dishonesty Score				0.082	0.020	0.167
Female	-0.034	-0.043	-0.061	0.006	0.031	-0.063
Framing	0.084	0.123	0.042	0.051	0.181	-0.076
STEM	0.016	0.117	-0.012	0.005	-0.050	0.111

Figure 4.5: Spearman Correlation Coefficients Between the Independent Variables and the Action of Choosing Control / Cheat (s1)

In the correlation analysis for $s1$, the D-factor consistently demonstrated a positive association across all game types, similar to the trend observed with the dishonesty score. Framing also correlates with strategy choices, particularly in one-shot games, with a stronger positive correlation for inspectors and a negative one for students. The "Female" variable showed a slight negative correlation with $s1$, suggesting that female participants might be less inclined towards these strategies. Age correlations were inconsistent, and the impact of STEM was minimal and context-dependent.

In the repeated game, we added two time-based variables: one reflecting the opponent's choice of control or cheat in the previous round (Opponent $S1$ (t-1)), indicating patterns like retaliation or imitation, and another tracking the player's historical cheat frequency (Own $S1$ Freq. (t-1)), representing their game "track record," included in the subsequent analyses. We then constructed logistic regression models for each dataset split to understand these relationships further. In these models, $s1$ was the dependent variable. It is important to note that in the repeated game, we treated every decision made by each player in each round as a distinct observation. Therefore, we clustered the standard errors at the subject level to account for repeated decisions by the same player across rounds. This method addresses the intra-participant correlation, considering potential impacts from unobserved individual traits or shared experiences, as highlighted by Bertrand, Duflo, and Mullainathan (2004). Although some independent variables exhibit a degree of correlation, we chose not to include interaction terms in the regression models. This decision was based on the absence of multicollinearity issues, as evidenced by the Variance Inflation Factors (presented in table 4.9, appendix 4.7.1). The results for the six models are summarized in table 4.2.

	Repeated Game			One-Shot Game		
	Game	P1(Inspector)	P2 (Student)	Game	P1(Inspector)	P2 (Student)
D-factor	0.51^{**}	0.51^{*}	0.54	0.44	0.01	1.17^{**}
Age	−0.15	−0.08	−0.39	0.06	0.04	0.28
Framing	0.14	0.25	−0.12	0.19	0.73^{**}	−0.46
Female	0.01	0.03	−0.22	0.11	0.09	0.09
STEM	−0.17	−0.18	−0.15	−0.03	-0.13	0.46
Dishonesty Score				1.19	0.31	3.22^{*}
Opponent S1 (t-1)	0.24	1.31^{***}	−0.36			
Own S1 Freq. (t-1)	4.78^{***}	4.38^{***}	5.12^{***}			
Constant	-3.67^{***}	-3.74^{***}	-2.32^{***}	-1.96^{**}	-0.26	-5.5^{***}
N	7000	3500	3500	526	264	262
Pseudo R-squared	0.29	0.28	0.16	0.01	0.03	0.09

Note: Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 4.2: Logistic Regression Results

In the repeated game, the D-factor significantly influenced the overall game and Player 1 in choosing $s1$, not Player 2. The opponent's previous choice of strategy 1 (Opponent $s1$ t-1) and the player's past behavior (Own $s1$ Freq. t-1) are highly significant for both players, indicating that past choices strongly guide current decisions. The relatively high Pseudo R-squared values suggest a good model fit, especially for Player 1. For the one-shot game, the D-factor was a significant predictor only for Player 2. The game's framing notably influenced Player 1's choice, while a higher Dishonesty Score significantly impacted Player 2's strategy selection. Although the Pseudo R-squared values are low, suggesting a weaker model fit than the repeated game, it is essential to note that the one-shot dataset has significantly fewer observations, rendering modeling tasks more challenging. Generally speaking, Age, Gender, and STEM fields did not significantly influence strategy choices in either game. The consistently negative and significant constants, particularly in the repeated game, suggest a baseline tendency against $s1$ when all other variables are at zero.

Overall, individual traits like the D-factor and past behavior have a stronger influence on the repeated game. In contrast, situational factors like framing and dishonesty score are more impactful in the one-shot game.

4.5.4 Retaliation and Reciprocity

We investigated the dynamics of reciprocity and retaliation to understand the relationship between individual characteristics and in-game actions within the repeated game context. Specifically, we looked at instances where players changed their strategies in response to their opponent's actions in the preceding round, particularly transitions from non-control to control or non-cheat to cheat. This analysis was further segmented by treatment and gender. The results are presented in table 4.3.

Comparison Dimension	Split	P1 (Inspector)	P2 (Student)
Treatment	Unframed	0.60	0.15
	Framed	0.65	0.17
Gender	Male	0.61	0.20
	Female	0.64	0.11

Table 4.3: Retaliation Frequencies Comparison

Inspectors retaliate more frequently than students, possibly due to the nature of their role, especially when facing cheating opponents. The Framed treatment shows higher retaliation rates for both roles than the Unframed treatment, indicating framing's impact on strategy. Among Inspectors, females retaliate slightly more than males, while in the Student role, males retaliate more than females. Using non-parametric tests like the Mann-Whitney U and Permutation tests (referenced in chapter 4.5.2), we found significant differences in retaliation based on player type but not for Gender or Framing, highlighting the strong influence of player type on retaliation. The full results are in table 4.8, appendix 4.7.1.

We broadened this analysis by incorporating additional variables into a logistic regression model, where the binary action of retaliation was the dependent variable. The standard errors in this model were also clustered at the participant level. The findings from this approach are detailed in table 4.4.

	P1(Inspector)	P2 (Student)
D-factor	-0.15	1.27^{**}
Age	0.13	-0.70^{*}
Framing	0.35	-0.27
Female	0.61	-0.74^{*}
STEM	0.37	-0.23
Own S1 Freq. (t-1)	2.03^{***}	5.62^{***}
Constant	-4.25^{**}	-4.67^{***}
N	3500	3500
Pseudo R-squared	0.06	0.24

Note: Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 4.4: Logistic Regression Results - Retaliation in the Repeated Game

The analysis of the regression model revealed distinct patterns of retaliation behavior. For Player 1, the past actions significantly predict their subsequent behavior rather than anything else. On the other hand, Player 2's decision-making is influenced not only by their past behavior but also by the D-factor, Age, and Gender. The model for Player 2 offers a more nuanced understanding, as reflected in the higher Pseudo R-squared value, highlighting different determinants for the distinct player types.

4.5.5 Random Forest Analysis

Following the same approach as the regression analysis in Chapter 4.5.3, we employed individual machine learning models for the overall game and each player. For this analysis phase, we utilized the Random Forest algorithm to better understand the variables influencing strategy choices across different game settings. These models' construction, training, and deployment adhered to the methodology outlined in Chapter 4.4. We calculated the SHAP values for each scenario using the training sets to interpret the outcomes. These values quantify the contribution of each independent variable to the prediction of strategy choices during the training process. The consolidated findings are presented in Figure 4.6.

In the repeated game scenario, past behavior is the most salient factor influencing both players' strategy selection, in accordance with the previous models. For Player 1, this tendency is further reinforced by the opponent's past choices, suggesting a more adaptive approach. On the other hand, Player 2 appears to be more self-reliant in their strategy selection but also shows a tendency to counterbalance their opponent's frequent choice of $s1$. While the D-factor moderately affects both players, its influence is also secondary to these behavioral patterns. The game's framing introduces a modest effect, subtly shaping players' strategic choices. Overall, the data underscores game behavior dominance in selecting strategies with a minor impact on personal features.

In the one-shot scenario, the features have a more subdued impact on strategy selection than the repeated game. Dishonesty Scores and D-Factor emerged as the most influential

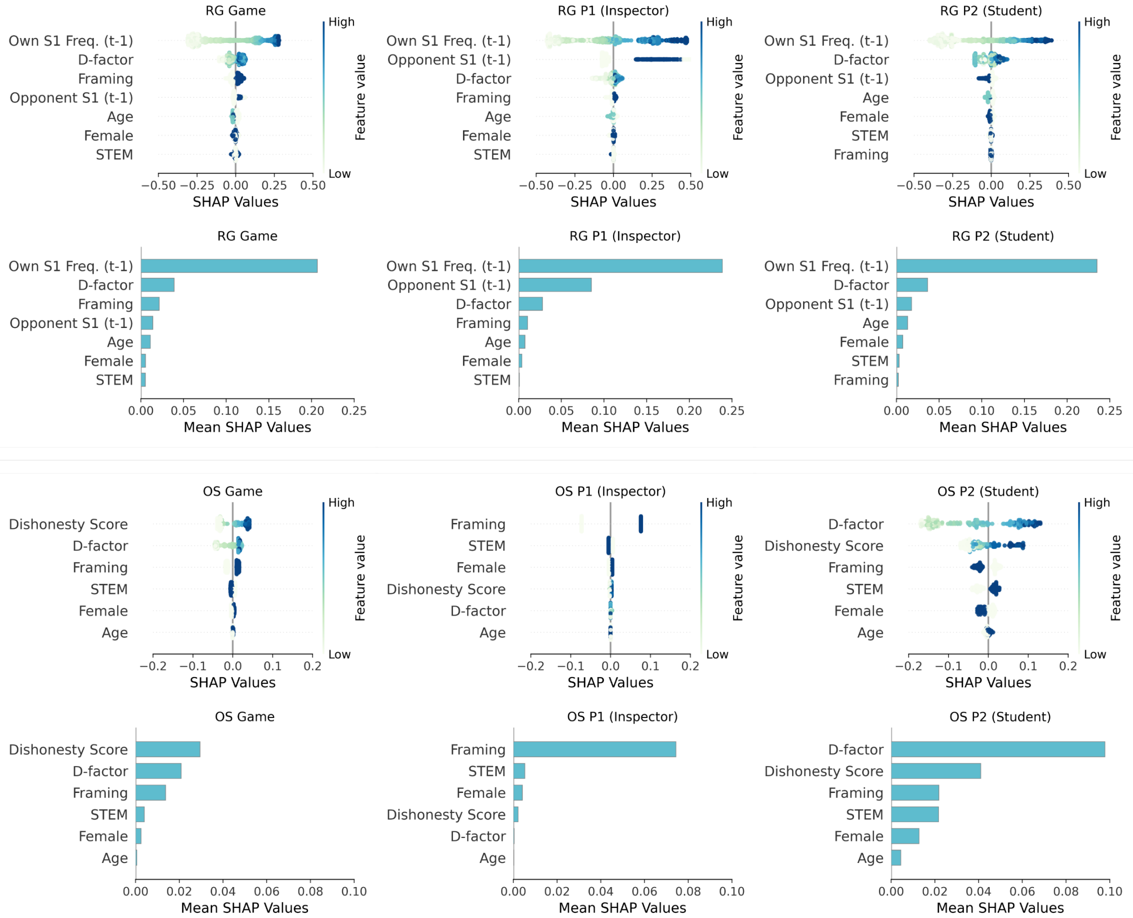


Figure 4.6: SHAP Values on the Traning Datasets - Random Forest

for the overall game, polarizing their effects at both ends. This distribution suggests that extreme values of the D-Factor are more decisive in strategy selection influence. Framing also plays a role but is more clearly divided: Those in the framed treatment were likelier to choose $s1$, while the unframed treatment was inclined to opt for the opposite. For Player 1, the influence of framing is even more pronounced, reinforcing that context could affect this player's choices. Player 2's choices are most influenced by high values of Dishonesty Scores and the D-Factor relative to the other variables. Framing produced the opposite effect on the student framing in this case, that is, driving less cheating.

The measurement of the models' feature importances (displayed in figure 4.11, appendix 4.7.1) generally reinforces the findings from the SHAP values. Regarding model performances, we computed the same set of performance metrics for all models, consolidated in table 4.5.

For the repeated game, all metrics are generally higher, indicating better model performance across the board. In contrast, the one-shot game shows notably lower values, especially for the game as a whole and Player 2 (Student), suggesting that the model struggles more with these scenarios. The drop in performance metrics for the one-shot game compared to the repeated game could indicate that the one-shot context is more

Dataset	Split	F1 Score	Precision	Recall	Accuracy	ROC-AUC
Repeated	Game	0.66	0.56	0.82	0.78	0.88
	P1 (Inspector)	0.73	0.73	0.74	0.78	0.87
	P2 (Student)	0.69	0.54	0.96	0.65	0.82
One-Shot	Game	0.43	0.40	0.47	0.50	0.48
	P1 (Inspector)	0.66	0.72	0.60	0.58	0.55
	P2 (Student)	0.25	0.20	0.33	0.66	0.59

Table 4.5: Random Forest Classifier Performance Metrics

challenging for the model to predict accurately due to the smaller dataset and the lack of patterns emerging from repeated interaction. In complement, the ROC curves for all models are plotted in 4.12, appendix 4.7.1.

In summary, the Random Forest analysis reveals that players' past behavior is again an influential factor in strategy selection in repeated games, especially for the inspector profile. Secondary variables such as the D-factor still pose a relevant influence but at a lower intensity. In one-shot games, the D-Factor and Framing variables become more prominent but have a subdued impact compared to repeated games. The model's performance metrics and feature importances corroborate these findings, highlighting the robustness of the behavioral patterns in the repeated game and the nuanced influences in the one-shot scenario.

4.5.6 Causal Inference Analysis and Treatment Effects

This section employs the Causal Forest Double Machine Learning (DML) model to disentangle covariate effects from treatment effects. This analytical separation allows for a detailed examination of participants' behavior in their strategic interactions, considering contextual (treatment) and individual factors. The first-stage models were selected using the procedure described in chapter 4.4. In most cases, the MLP Classifier was the model of choice for T and Y, consistently outperforming other models. The only exception was the Y model in the one-shot game dataset, where a Random Forest Classifier was selected instead. The overview of the model selection results is documented in table 4.11, appendix 4.7.3. The frequent selection of the Neural Network-based model across all scenarios suggests that its complexity and flexibility make it better suited for capturing the intricate patterns in the data, as evidenced by the higher overall scores. In addition, the selected hyperparameters for each model instance are documented in table 4.12.

In constructing the Causal Forest DML models, we followed a similar approach to the Random Forest models. The dependent variable (Y) is the strategy choice '$s1$,' while the personal attributes serve as covariates (X), and 'Framing' is designated as the treatment variable (T). By accounting for past decisions, we ensured that our observed impact was genuinely due to 'Framing' and not influenced by previous game dynamics. Therefore, the time-based variables were defined as cofounders (W). We computed the SHAP values to interpret the models' results, displayed in figure 4.7. The SHAP values in this context measure the independent variables' influence on the framing effects' intensity, not in the decisions directly.

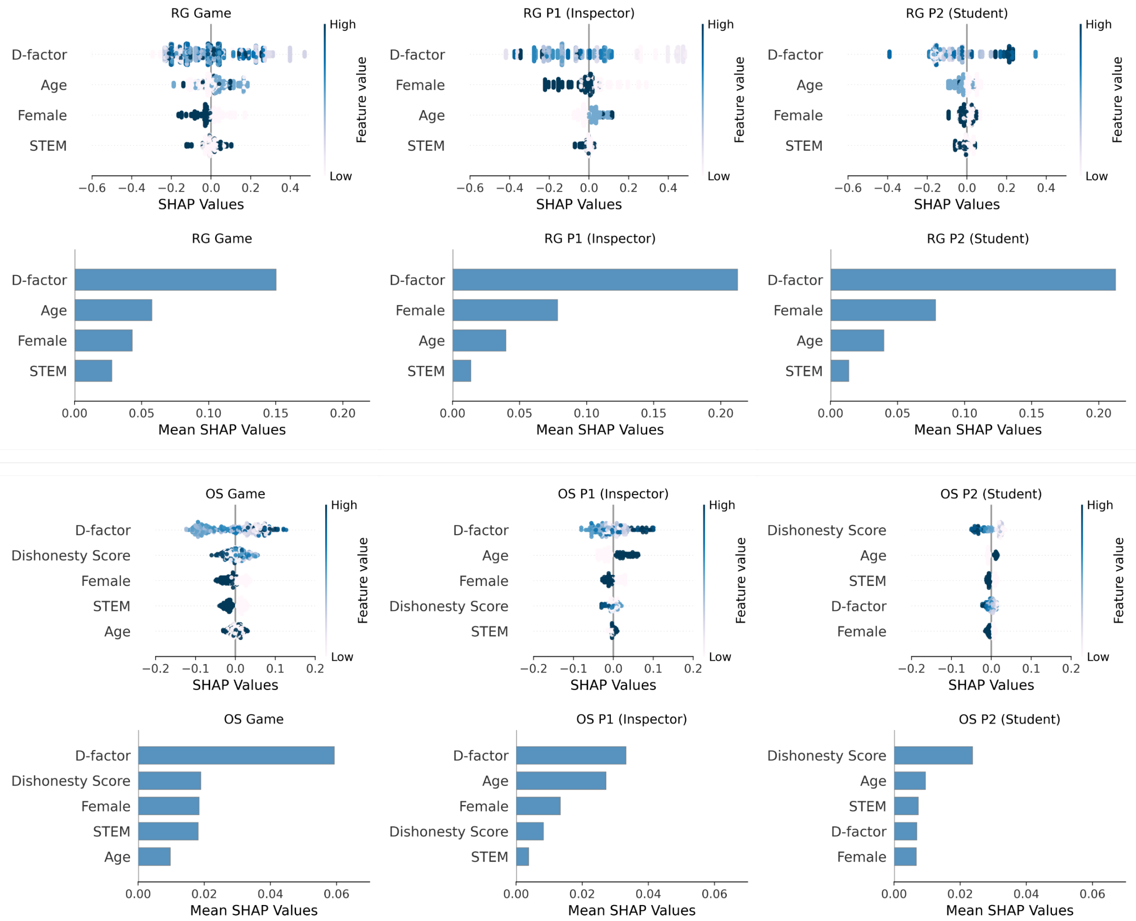


Figure 4.7: SHAP Values on the Training Datasets - Causal Forest (DML)

In the Repeated Game, the D factor stood out as a significant influence. Younger players tended to be negatively affected by the game's framing, while older ones experienced positive effects. Gender differences were evident, with females being more negatively influenced than males. For Player 1, those with low D-factor values were particularly affected, whereas younger Player 2s showed positive effects. In the one-shot game, the D factor's influence was less dominant but still notable. The same gender patterns persisted, with females generally being more negatively affected. High D factor values led to more positive outcomes for Player 1, while Player 2's decisions were more influenced by their dishonesty scores. Across both games, the D factor, age, and gender consistently impacted the effects of framing, but their influence varied depending on the game type and player role.

We assessed the Conditional Average Treatment Effects (CATE) using causal forest models to understand treatment impact conditioned on covariates. The effect estimations on the training sets provide an average view of the treatment's influence. In Figure 4.8, the upper plots show the CATE for individual observations, with a 95% confidence interval shaded, while the lower histograms present the frequency distribution of these effects. These visualizations offer a clear comparison of treatment effects across the different dataset splits.

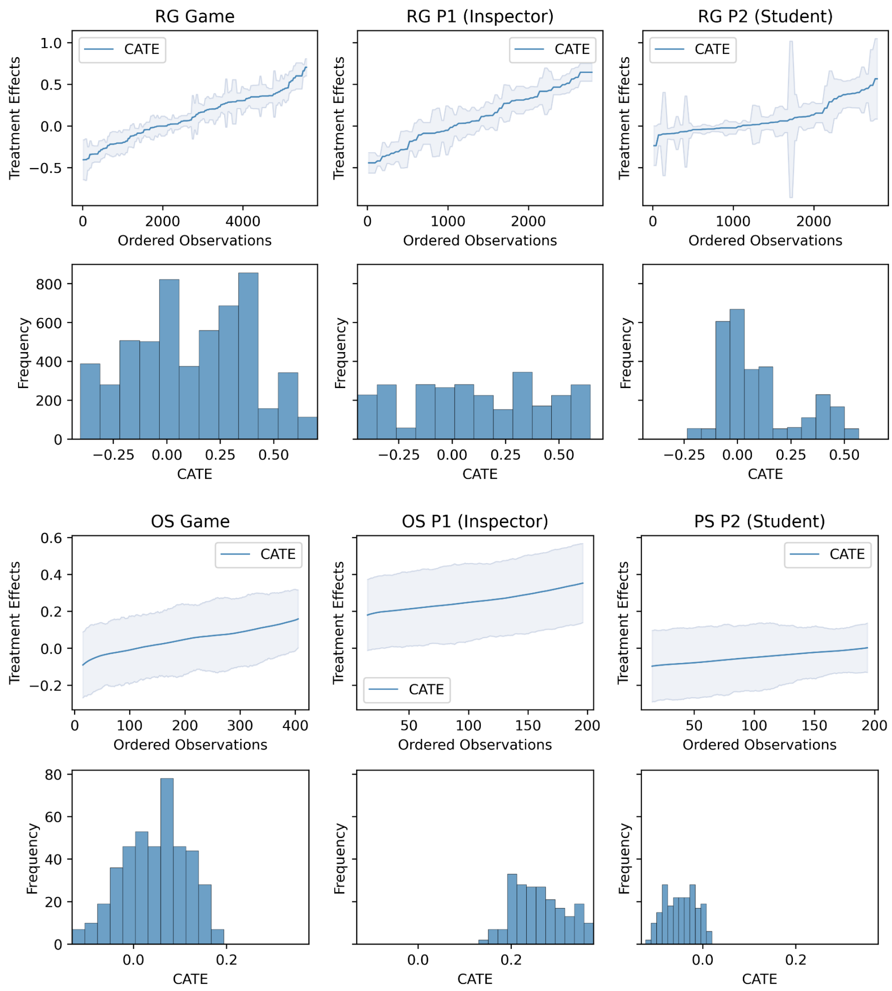


Figure 4.8: CATE on the Training Datasets - Causal Forest (DML)

There is a noticeable trend in the analysis of repeated games: the distribution of effects is broader, with the inspector players exhibiting higher effects than the student players. This pattern is even more pronounced in the one-shot dataset. Here, most positive effects are attributed to the inspector player, while negative effects are predominantly associated with the student player. The treatment effect trends suggest a relatively higher influence on individuals fitting the inspector profile.

To validate the robustness of our effect quantifications, we turn to predictions on the test set, evaluating the model's performance. We employ a widely recognized metric from foundational studies by (Rosenbaum & Rubin, 1983; Rubin, 1974), the Average Treatment Effect (ATE)[2]. This metric captures the anticipated outcome difference between the treated and untreated groups across the entire dataset, offering a comprehensive perspective on the overall impact of the treatment. The outcomes of these predictions are detailed in table 4.6.

The treatment effect predictions point in the direction that the framing was generally more

[2]The ATE is expressed as $E_X[\tau(X, T0, T1)]$

Dataset	Split	ATE	MAE	RMSE
Repeated	Game	0.107	0.532	0.393
	P1 (Inspector)	0.112	0.635	0.516
	P2 (Student)	0.091	0.358	0.226
One-Shot	Game	0.039	0.556	0.365
	P1 (Inspector)	0.238	0.540	0.470
	P2 (Student)	−0.035	0.410	0.189

Table 4.6: Predicted Average Treatment Effects (ATE) and Error Measures on the Test Datasets

influential for the Player 1 group, as indicated by higher ATE values in both ”Repeated” and ”One-Shot” datasets (consistent with earlier findings). However, this group also exhibited higher error measures, suggesting less reliable predictions. In contrast, the second player's group had the most reliable predictions in the ”Repeated” dataset. However, it showed a negative ATE in the ”One-Shot” dataset, indicating that the framing might have prevented cheating for this group. The game split has moderate ATE values. The predicted effects generally align with the observed participants' behavior in the game.

The consistent prominence of D-factor, age, and gender across game scenarios underscores their significance in influencing treatment effects. Moreover, the differential impact of the treatment on inspector and student profiles, as evidenced by the ATE values, offers valuable insights into the nuanced dynamics of the game. While the predictions validate the general trends observed, they also highlight areas of uncertainty, particularly for the inspector group. This detailed analysis provides insights into the framing effects not identified earlier with regular statistics.

4.6 Conclusion

This study analyzed the influence of dark personality traits within an inspector game, comparing an information-loaded frame to a neutral one. The combination of game mechanics, contextual changes, behavioral tendencies, and individual attributes presents a complex environment for strategic decision-making. Our findings provide a comprehensive understanding of strategic decision-making, emphasizing the distinct roles of individual attributes and treatment effects in shaping players' behaviors in the game.

Participants often deviated from the Nash Equilibrium's theoretical predictions, favoring less risky and morally neutral strategies. These choices often resulted in more favorable outcomes, highlighting the nuanced dynamics of the game. Notably, the inspector profile demonstrated a heightened sensitivity to perceived dishonesty, especially under the framed conditions.

The effects of framing, while subtle, were evident. When informed about the student's motivations, the inspector profile showed an increased tendency to control more in the framed treatment. However, when personal attributes were integrated into the analysis, the dominant role of game mechanics became clear. The D-factor, although influential, was secondary to the rational behaviors aimed at maximizing outcomes. This dominance was further emphasized in advanced models like the Random Forest.

The causal forest model provided insights into the relationship between treatment effects and individual attributes. The inspector profile consistently showed more pronounced effects, with the dark factor emerging as a significant predictor of frame effects. The combined use of the causal forest and SHAP values effectively captured these intricate effects.

Models derived from the repeated game dataset demonstrated more robust performance, likely due to the iterative nature of these games. In contrast, the one-shot dataset showed less reliable and consistent performance. As such, predictions from the latter models should be approached with caution.

This study has limitations, including the complexity of the experimental construct, potential biases, and unobserved confounders. The computational demands of models like DML also present challenges. The specificity of the game in this study suggests an avenue for future research, exploring a variety of games beyond moral dilemmas. We also understand that this is a specific game with very particular mechanics.

In conclusion, this research highlights the utility of machine learning methods in understanding the complex behavioral dynamics of strategic games. By elucidating the interactions between individual attributes, game mechanics, and contextual influences, this study contributes to a broader understanding of strategic decision-making.

4.7 Appendix

4.7.1 Additional Data Analysis Elements

The results of the employed statistical tests for the decision-based samples are documented in table 4.7. The direction is positive in all cases, which means Sample A > B. Similarly, the test results for the retaliation samples are in table 4.8, and the directions are also positive.

Dataset	Sample A	Sample B	Test Results
Repeated	Framed	Unframed	Mann-Whitney U: 0.614, Permutation: 0.092
Repeated	Male	Female	Mann-Whitney U: 0.904, Permutation: 0.291
Repeated	P1 Framed	P1 Unframed	Mann-Whitney U: 0.256, Permutation: 0.083
Repeated	P2 Framed	P2 Unframed	Mann-Whitney U: 0.946, Permutation: 0.253
One-Shot	Framed	Unframed	Permutation: 0.134, Z-test: 0.239, Fisher's: 0.242
One-Shot	Male	Female	Permutation: 0.325, Z-test: 0.578, Fisher's: 0.586
One-Shot	**P1 Framed**	**P1 Unframed**	**Permutation:** 0.002, **Z-test:** 0.003, **Fisher's:** 0.004
One-Shot	P2 Framed	P2 Unframed	Permutation: 0.917, Z-test: 0.220, Fisher's: 0.263

Table 4.7: Game Decisions Statistical Tests Overview - p Values

The distribution of D scores for both samples is plotted in figure 4.9. The one-shot sample, which is larger, provides a more realistic distribution of the D scores, according to the

Sample A	Sample B	Test Results
Framed	Unframed	Mann-Whitney U: 0.342, Permutation: 0.661
Male	Female	Mann-Whitney U: 0.869, Permutation: 0.656
Inspector	Student	Mann-Whitney U: 0.001, Permutation: 0.000

Table 4.8: Retaliation Statistical Tests Overview (Repeated Game) - p Values

overall distributions documented on the D-factor questionnaire website[3].

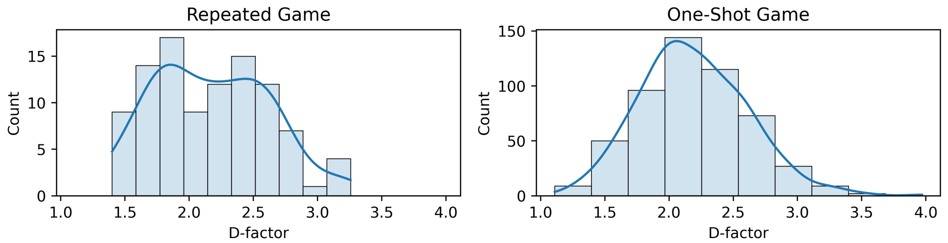


Figure 4.9: Distributions of D-Factor Scores, 1 is the Lowest and 5 is the Highest

Figure 4.10 displays the Spearman correlation coefficients for all variables in both game datasets.

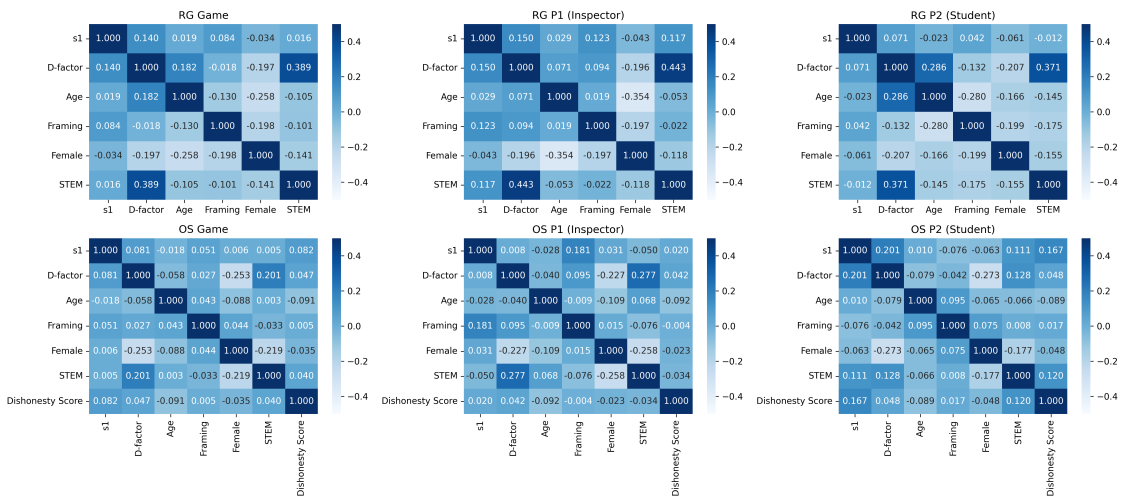


Figure 4.10: Spearman Correlation Coefficients for the Repeated (RG) and One-Shot (OS) Samples

In order to understand the relationships between our independent variables, we computed the Variance Inflation Scores (VIF), a method used to detect multicollinearity in regression models (Draper & Smith, 1998). The results are compiled in table 4.9.

In both the Repeated and One-Shot Game datasets, the Variance Inflation Factors (VIF) for all independent variables range from 1.01 to 1.33, well below the commonly cited

[3] https://www.darkfactor.org/

Variable	Repeated Game	One-Shot Game
D-factor	1.33	1.10
Age	1.21	1.03
Framing	1.13	1.01
Female	1.18	1.11
STEM	1.33	1.07
Dishonesty Score		1.02
Opponent S1 (t-1)	1.03	
Own S1 Freq. (t-1)	1.04	

Table 4.9: Variance Inflation Factors (VIF) Within Independent Variables

threshold of 5. This suggests that multicollinearity is unlikely to be a significant concern in the regression models for either dataset.

Feature importances provide a global measure of each variable's impact on the model's predictions. They differ from SHAP values because they offer an aggregate and simplified view rather than instance-level explanations. Figure 4.11 displays the computed feature importance for the Random Forest model.

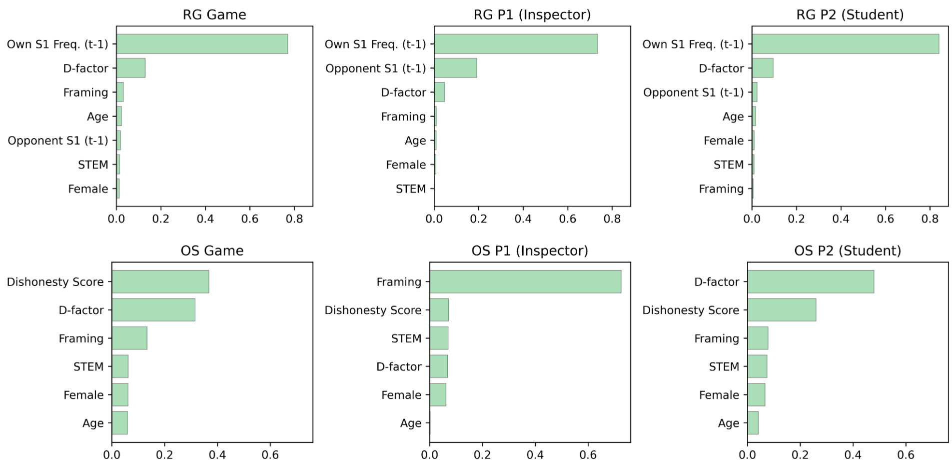


Figure 4.11: Feature Importances - Random Forest

In the Repeated Game, "Own $s1$ Freq. (t-1)" is the most important feature for all models, followed by "D-factor." In the One-Shot Game, "Dishonesty Score" and "D-factor" dominate the game model, while "Framing" is significantly important for Player 1. For Player 2 in the One-Shot Game, "D-factor" takes precedence. The importance of features varies significantly between the repeated and one-shot games and between players within each game type.

The ROC curves evaluate the performance of a classification model by plotting the true positive rate against the false positive rate (as described in chapter 4.4). AUC (Area Under the Curve) values range from 0 to 1, with higher values indicating better model performance. The ROC plots for the random forest model are compiled in figure 4.12.

In the repeated game, the models for the game, Player 1 and Player 2, show strong predic-

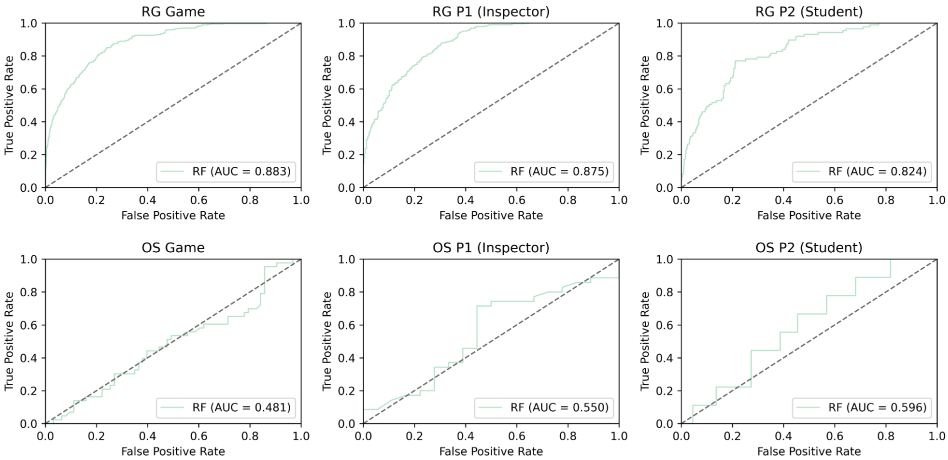


Figure 4.12: AUC-ROC Curves - Random Forest

tive power with AUC values above 0.8. In contrast, the one-shot game model performance is less satisfactory, confirming a difficult prediction scenario. This is mainly due to the dataset sizes and difficulty finding clear trends in a one-shot framework.

4.7.2 Theoretical Formalizations

Cyclic Games and Nash Equilibrium

The basic taxonomy of cyclic games in matrix form, which includes the inspector game, is given by Selten and Chmura (2008). A matrix will be a cyclic game if the following conditions are met (table 4.10):

		Pl 2	
		L	**R**
Pl 1	**U**	$(\mathrm{a_L + c_L, b_U})$	$(\mathrm{a_R, b_U + d_U})$
	D	$(\mathrm{a_L, b_D + d_D})$	$(\mathrm{a_R + c_R, b_D})$

$a_L, a_R, b_U, b_D \geq 0$; $c_L, c_R, d_U, d_D > 0$

Table 4.10: Taxonomy of experimental 2×2 cyclic games

The Nash Equilibrium is a strategy profile in which every strategy is an optimal response to other players' strategies (Nash Jr, 1950). The basic NE formula for computing the equilibrium of cyclic games is given as follows:

$$p_U = \frac{d_D}{d_U + d_D}, \quad q_L = \frac{c_R}{c_L + c_R}. \tag{4.3}$$

For more details in stationary equilibria computation for cyclic games, see Brunner et al. (2010, 2011); Selten and Chmura (2008).

Dishonesty Scores

We utilized a randomized response strategy with a coin flip, where the flip's outcome determined if a participant should answer five dishonesty-related questions truthfully or randomly. Given this uncertainty, we estimated the true proportion of "yes" responses for each question by:

$$\text{Proportion of "yes"} = \frac{\text{Observed proportion of "yes"} - \text{probability of random "yes"}}{\text{Probability of truthful response}} \tag{4.4}$$

For a fair coin flip dictating truthful or random responses, this formula simplifies to:

$$\text{Proportion of "yes"} = 2 \times (\text{Observed proportion of "yes"} - 0.25) \tag{4.5}$$

We then aggregated the estimated proportions from all questions to derive a single "dishonesty score", averaged across questions. This score ranges between 0 and 1, where higher values suggest a greater likelihood of dishonest behavior. For participants answering based on the coin flip, the method presumes a 0.5 probability of getting a 'yes' due to the coin flip alone. Hence, half of the 'yes' responses could be attributed to the coin, and the other half to actual affirmative responses. The true 'yes' proportion is estimated by subtracting the 0.25 probability (from the coin flip) from observed proportions and doubling the remaining amount. If the resultant number is negative, it's adjusted to 0.

Random Forest

The Random Forest algorithm concept builds a large collection of de-correlated decision trees and then aggregates them through a majority voting system for classification problems. A generalization of the implementation of a random forest classifier is available in Hastie, Tibshirani, Friedman, and Friedman (2009). The algorithm can be generalized in algorithm 1.

Algorithm 1 Random Forest Algorithm (Generalized)

Require: B trees to be grown, N size of bootstrap sample, M total variables, m selected variables, n_{min} minimum node size
Ensure: Output the ensemble of trees $\{T_b\}_1^B$

for $b = 1$ to B **do**
    Draw a bootstrap sample of size N from the training data
    Grow a decision tree T_b on this data by:
    while each terminal node of the tree until the minimum node size n_{min} is reached **do**
        Select m variables at random from all M variables
        Pick the best variable/split-point among the m
        Split the node into two daughter nodes
    end while
end for
To make a prediction for a new point x, let $\hat{C}_b(x)$ be the class prediction of the bth random forest tree
The random forest chooses $\hat{C}_{\text{rf}}(x) = \text{majority vote}\{\hat{C}_b(x)\}_1^B$

Further information about the model is provided in Breiman (2001).

Causal Forest and Double Machine Learning Estimators

The transition from a Random Forest to a Causal Forest primarily focuses on the target variable and the splitting criterion of the tree. In Random Forests, the trees aim to minimize classification or regression errors (classification, in our case), but in Causal Forests, they are designed to estimate causal effects. Double Machine Learning (DML) is a method that integrates decision tree ensembles with the DML approach to estimate localized causal treatment effects robustly. This method is designed to estimate treatment effects in the presence of high-dimensional covariates while ensuring that confounders are properly accounted for. The method is called "double" because it typically involves two sets of machine learning predictions. The first model (model T) predicts the treatment T given covariates X. This provides a propensity score or probability of receiving the treatment based on observed characteristics. The second model (model Y) predicts the outcome Y given covariates X. This gives an idea of the expected outcome based solely on observed characteristics without considering the treatment. Essentially, model T is trying to approximate the conditional expectation $f(X, W) = \mathbb{E}[T|X, W]$, and model Y is capturing $g(X, W) = \mathbb{E}[Y|X, W]$. These are non-parametric regression tasks, where we're trying to predict the expected value of our treatment or outcome based solely on observed characteristics.

Using the predictions from the two models, DML computes the residuals $\tilde{T}$ and $\tilde{Y}$ by subtracting predicted values from the observed treatment and outcome, respectively. These residuals represent the component of treatment and outcome that the observed covariates cannot explain.

This causal forest variant, grounded in the principles of Double Machine Learning, aims to estimate causal effects by utilizing a blend of machine learning techniques and econometric insights. In particular, the CausalForest methodology described here is based on a residual-on-residual local moment condition:

$$\hat{\theta}(x) = \arg\min_{\theta} \sum_{i=1}^{n} K_x(X_i) \cdot \left(Y_i - \hat{q}(X_i, W_i) - \theta \cdot (T_i - \hat{f}(X_i, W_i))\right)^2 \tag{4.6}$$

X_i represents the set of observed covariates or features for the i^{th} observation that are directly of interest in assessing the treatment effect. W_i denotes additional covariates or features for the i^{th} observation, which, while not the primary focus, can influence the treatment and outcome. The goal is to find the estimate $\hat{\theta}(x)$ that minimizes the squared difference between the observed outcome Y_i and its predicted value after adjusting for the treatment effect and its prediction. This is done in a 'local' fashion using a weight function $K_x(X_i)$ that gives more importance to data points close to x. $K_x(X_i)$ represents a similarity metric, capturing how similar observation i is to the target point x. $\hat{q}(X_i, W_i)$ represents the predicted outcome value based on covariates X_i and W_i. $\hat{f}(X_i, W_i)$ gives the predicted treatment value based on the same covariates. The model considers two distinct splitting criteria. The first is based on the mean-squared error (MSE):

$$\max_{S_1,S_2} \theta_1^2 \sum_{i\in S_1} \tilde{T}_i^2 + \theta_2^2 \sum_{i\in S_2} \tilde{T}_i^2 \approx \max_{S_1,S_2} \theta_1^2 \cdot |S_1| \cdot Var_n(T|x \in S_1) + \theta_2^2 \cdot |S_2| \cdot Var_n(T|x \in S_2) \quad (4.7)$$

Which partitions the data into two subsets, S_1 and S_2, such that the sum of squares of the treatment effect estimates θ_1 and θ_2 on these subsets is maximized. This is done by factoring in the variability (as denoted by $\tilde{T}_i$ within each subset. The criterion prioritizes maximizing heterogeneity while penalizing splits, leading to low treatment variability within nodes.

The second criterion is the one suggested in Athey and Wager (2019). Unlike the MSE Criterion, Athey's Criterion (HET) aims to maximize the treatment effect's heterogeneity without considering the treatment's within-node variability. It solely optimizes for maximizing the distinctiveness between the two estimates.

$$\max_{S_1,S_2} \theta_1^2 + \theta_2^2 \quad (4.8)$$

Based on Battocchi et al. (2023)'s implementation of Chernozhukov et al. (2017); given a dataset with treatments T, outcomes Y, and potential confounders X, the overarching objective is to discern the causal effect of T on Y conditional on X. The model postulates the following structural relationships for the data-generating process:

$$\begin{aligned} Y &= \theta(X) \cdot T + g(X,W) + \epsilon, \qquad &\mathbb{E}[\epsilon|X,W] = 0 \\ T &= f(X,W) + \eta, &\mathbb{E}[\eta \mid X,W] = 0 \\ &\mathbb{E}[\eta \cdot \epsilon|X,W] = 0 \end{aligned} \quad (4.9)$$

In the equations in 4.9, Y represents the outcome, influenced by the product of the treatment effect $\theta(X).T$, a function g that encapsulates the effects of covariates X and cofounders W, and an error term ϵ. T stands for the treatment, defined by a function f and an error term η. The errors ϵ and η are assumed to be uncorrelated when conditioned on the covariates X and cofounders W.

To estimate the constant marginal Conditional Average Treatment Effect (CATE), namely $\theta(X)$, distilling the pure effect of the treatment from the confounding influences, the outcome Y is residualized concerning its expected value, leading to:

$$Y - \mathbb{E}[Y|X,W] = \theta(X).(T - \mathbb{E}[T|X,W]) + \epsilon \quad (4.10)$$

For estimation, we consider the conditional expectation functions as non-parametric regression tasks:

$$\begin{aligned} g(X,W) &= \mathbb{E}[Y|X,W] \\ f(X,W) &= \mathbb{E}[T|X,W] \end{aligned} \quad (4.11)$$

From the regressions, we compute the residuals, representing the unexplained variances of Y and T as:

$$\begin{aligned} \tilde{Y} &= Y - q(X, W) \\ \tilde{T} &= T - f(X, W) = \eta \end{aligned} \tag{4.12}$$

These residuals provide a relationship governed by:

$$\tilde{Y} = \theta(X).\tilde{T} + \epsilon \tag{4.13}$$

Owing to the condition $\mathbb{E}[\epsilon.\eta|X] = 0$, the treatment effects $\theta(X)$ can be estimated by regressing the residual outcome $\tilde{Y}$ on the residual treatment $X, \tilde{T}$ and covariates X, for example:

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \mathbb{E}_n[(\tilde{Y} - \theta(X).\tilde{T})^2] \tag{4.14}$$

By adhering to these structured steps, the Double Machine Learning approach isolates and quantifies the treatment effect, filtering out confounding influences and capturing the pure causal impact of the treatment.

Shapley Values

Considering a machine learning model as a real-valued function f that takes a vector of real-valued features as input. For a classification problem, the function models the score of a class. The set of model features is denoted by N. The vector x of features denotes the input to be explained or the explicand. x_s expresses the sub-vector of a vector x restricted to the features in the set S. Formally, a cooperative game with a characteristic function v: $2^N \to R$ (considered as the models's input), where N is the set of players, player i's Shapley value s_i is given as:

$$s_i = \sum_{S \subseteq N \setminus i} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [v(S \cup i) - v(S)] \tag{4.15}$$

This formula sums over all possible coalitions S that do not include player i, and distributes the payoffs according to the marginal contribution of player i to each coalition. The term $\frac{|S|!(|N|-|S|-1)!}{|N|!}$ refers to the probability that coalition S forms before player i joins when other players join one at a time, randomly (Shapley et al., 1953). The Shapley value can also be described using a permutation approach: Players are randomly ordered, sequentially added, and each player i is assigned its expected marginal contribution, given by $V(S \cup i) - v(S)$ where S consists of players added before i.

Following S. M. Lundberg and Lee (2017)' approach, in the context of machine learning models, the key application of the Shapley values is based on the concept's extension known as the Conditional Expectations Shapley (CES). This variation takes three inputs:

an explicand x, a function f, and a distribution D. The conditional expectation defines the set function:

$$v(S) = E_D[f(x')|x'_S = x_S] \tag{4.16}$$

The algorithmic procedure to compute the CES is given in algorithm 2.

Algorithm 2 Calculating the Shapley values with CES($\hat{D}$)

Require: explicand x and examples T, each over feature set N
Ensure: Compute Shapley values via permutations

Initialize $s_{\sigma i} \leftarrow 0$ for all i
for all permutations σ of N **do**
  $v_{\text{new}} \leftarrow \frac{1}{|T|} \sum_{x \in T} f(x)$
  $T' \leftarrow T$
  for all $i \in \{1, \ldots, |N|\}$ **do**
    $v_{\text{old}} \leftarrow v_{\text{new}}$
    Update T'
    for all $t \in T'$ **do**
      if $x^t_{\sigma i} \neq x_i$ **then**
        remove t from T'
      end if
    end for
    $v_{\text{new}} \leftarrow \frac{1}{|T'|} \sum_{x \in T'} f(x)$
    Update the Shapley value of the i-th feature ordered in σ
    $s_{\sigma i} \leftarrow s_{\sigma i} + \frac{1}{|N|!}(v_{\text{new}} - v_{\text{old}})$
  end for
end for

The CES provides an intuitive way to attribute the impact of each feature on a given prediction based on its actual value and the data distribution. For more details on this concept, including axiomatic considerations, see Sundararajan and Najmi (2020).

4.7.3 Technical Remarks

The experiment was programmed using oTree (D. L. Chen et al., 2016). The regression models and statistical tests used the statsmodels python library (Seabold & Perktold, 2010). The remaining machine learning models were implemented using Scikit-Learn in python (Pedregosa et al., 2011). The TPE hyperparameter estimation for the Random Forest analysis applied the optimization algorithms from Optuna (Akiba, Sano, Yanase, Ohta, & Koyama, 2019). The causal forest models used the EconML in python (Battocchi et al., 2023). The causal forest DML models were tuned using the native EconML tuning function.

Model Selection, Training, and Tuning Details

For the random forest, the models were tuned using the Tree-structured Parzen Estimator (TPE) algorithm for hyperparameter optimization, which uses the history of previously evaluated hyperparameter configurations to sample the following ones (Akiba et al., 2019).

The optimization maximized the ROC-AUC metric over 1000 trials for the six distinct cases (Akiba et al., 2019).

In the causal forests framework, we assessed the Random Forest, Gradient Boosting Machines, Gaussian Naive Bayes, and the Multi-layer Perceptron as candidates for models T and Y. The random forest model has been previously described. The remaining three models' descriptions are provided next.

Gradient Boosting Machines (GBMs) are ensemble learning methods that build predictive models by iteratively adding weak learners to minimize residual errors. The technique employs boosting to optimize a differentiable loss function, introduced in J. H. Friedman (2001). NB classifiers are probabilistic models that apply Bayes' theorem with strong feature independence assumptions. The Gaussian variant of NB assumes feature likelihoods are Gaussian-distributed (Bishop & Nasrabadi, 2006; Murphy, 2012). The Multi-Layer Perceptron (MLP) is an artificial neural network consisting of interconnected layers of nodes, or "neurons," each applying a non-linear activation function. Formulated initially in early work by Rumelhart, Hinton, and Williams (1986), MLPs are prominent for the ability to model complex, non-linear relationships in data.

We employed grid search techniques to select the best-fitting model among these four options systematically. This algorithmic approach exhaustively tests all possible parameter combinations within a predefined search space and is detailed further in Bishop and Nasrabadi (2006); Hastie et al. (2009). We gauged the performance of each model using composite scores of basic classification metrics (as mentioned earlier for the Random Forest), RMSE, and MSE. It is crucial to note that the computational complexity of grid search can be prohibitive, depending on the breadth of the search space. Upon completing the model selection, the DML model uses the selected models to generate CATE predictions.

Specific cross-validation methods were employed (Kohavi et al., 1995) to ensure the robustness and reliability of model predictions. In repeated games, we used group-level cross-validation per participant, preventing overfitting. For one-shot games, Stratified K-fold cross-validation ensured dependable estimates. Due to causal modeling's computational intensity, we used three-fold cross-validation, while random forest models employed a five-fold approach.

DML - First-Stage Models and Model Selection Procedure

A consistent methodology was applied across all datasets in estimating the first-stage models or the causal forest. However, special attention was required for the one-shot dataset due to its significantly smaller size. Theoretical considerations suggest that applying highly complex models to small datasets is prone to overfitting, which can subsequently distort the results of second-stage models, even when cross-validation techniques are employed.

We modified the model estimation function specifically for the one-shot dataset to mitigate this risk. This modification restricts the complexity of the models by limiting the architecture of neural networks and constraining the depth of trees in ensemble models.

In addition, we employ a higher number of cross-validation folds. This approach aims to balance model complexity and predictive power, thereby enhancing the reliability of the second-stage estimates. The scores achieved by each model for each dataset and the selected variants in each case are documented in table 4.11.

Dataset	Split	Model	MLP	RF	GBM	GNB	Model Selected
Repeated	Game	T	0.657	0.551	0.582	0.530	MLPClassifier
		Y	0.596	0.496	0.455	0.455	MLPClassifier
Repeated	P1 (Inspector)	T	0.680	0.526	0.426	0.476	MLPClassifier
		Y	0.642	0.511	0.486	0.536	MLPClassifier
Repeated	P2 (Student)	T	0.784	0.655	0.659	0.531	MLPClassifier
		Y	0.610	0.531	0.572	0.482	MLPClassifier
One-Shot	Game	T	0.564	0.524	0.532	0.463	MLPClassifier
		Y	0.529	0.544	0.536	0.495	RandomForestClassifier
One-Shot	P1 (Inspector)	T	0.584	0.536	0.558	0.508	MLPClassifier
		Y	0.571	0.527	0.538	0.513	MLPClassifier
One-Shot	P2 (Student)	T	0.597	0.540	0.550	0.517	MLPClassifier
		Y	0.630	0.627	0.607	0.556	MLPClassifier

Table 4.11: The first-stage models' performance, based on the ROC-AUC refit scores.

Each first-stage model was fit individually to its respective dataset systematically. The first-stage model selection procedure selected the models' architectures. The hyperparameter configuration for each model used is documented in table 4.12.

Random Forest and Causal Forest DML Parameters

The optimized parameters for the Random Forest models applied in chapter 4.5.5 are documented in table 4.13.

For the causal forest model, we established an initial set of parameters aligned with our research objectives. This foundational model was consistently applied to all samples. However, specific parameters were fine-tuned for each sample using a dedicated tuning function. The universally applied model parameters were set as follows: `n_estimators: 1000`, `inference: True`, `discrete_treatment: True`, `cv: 10`, `drate: True`. In this context, the inference parameter enables statistical inference on causal effects by computing standard errors when set to True. The CV parameter defines the cross-validation strategy, enhancing the robustness of the treatment effect estimates. Additionally, when activated, the drate parameter leverages the double robustness property, ensuring consistent treatment effect estimates if the treatment effect or the outcome model is correctly specified.

4.7.4 Design Elements

This section documents the experiment design applied to the participants. Figure 4.13 contains the explanation screen common to all participants in our samples. Figure 4.14 shows the framed version of the game, containing the contextual explanations and named strategies. Conversely, figure 4.15 displays the unframed version with generic strategies and no context.

Introduction - How matrix games work

We'll give a short introduction on how the games -called Matrix games- are played: Matrix games describe two person (Player 1 & Player 2) decision situations.

Example:

		Player 2	
		A	B
Player 1	A	(2, 2)	(0, 3)
	B	(3, 0)	(1, 1)

In this example, the possible strategies for each player are named A & B and each player can pick one of these two strategies

The corresponding payoffs are given in the brackets. The first number is the payoff for player 1 and the second number is the payoff for player 2. So for example, (5, 2) denotes a payoff of 5 for player 1 and 2 for player 2.

Therefore, the final payoff depends not only on the own individual decision, it is also based on the decision of the other player. But player 1 and 2 choose their strategies without knowing the decision of the other player.

Example: If player 1 chooses strategy A, and player 2 strategy B, we find this situation:

		Player 2	
		A	B
Player 1	A	(2, 2)	(0, 3)
	B	(3, 0)	(1, 1)

The payoff for player 1 is 0 and for player 2 it is 3.

You will play against another human player, who will be chosen randomly in the beginning, but will remain the same for all subsequent rounds.

Figure 4.13: Game explanation

Round 1 of 70

Explanation:

You are a student in an exam situation.

You can decide in each round whether "to cheat" or "not to cheat" on the examiner. You will be matched with a second player (the examiner) in each round, who can decide to "control" or "not control" you.

The payoffs for all four possible game outcomes are shown in the matrix.

Example: If you choose "Don't Cheat" and the other player chooses "Control", you will receive 10 experiential units and the other player 11 experimental units.

You receive the highest possible payoff of 16 experimental units if you chose "Cheat" ant the other player Chose "Don't Control".

You receive the lowest possible payoff of 2 experimental units if you chose "Cheat" and the other player chose "Control".

Choose one of the two strategies! Your opponent has already made a strategic choice.

Your choice for this round:

		Examiner	
		A	B
Student (you)	Cheat	(2, 12)	(16, 3)
	Don't Cheat	(10, 11)	(14, 14)

Figure 4.14: Framed design example

For the one-shot game, the following questions have been applied to compute the dishonesty scores:

1. In your studies, have you ever copied from other students during an exam?

Figure 4.15: Unframed design example

2. In your studies, have you ever used illicit crib notes in an exam?
3. In your studies, have you ever used prescription drugs to enhance your performance in an exam
4. In your studies, have you ever handed in a paper containing a passage intentionally adopted from someone else's work without citing the original?
5. In your studies, have you ever had someone else write a large part of a submitted paper for you, or have you handed in someone else's paper as your own?

The dishonesty questions also included the coin-based randomized response method. Figure 4.16 shows how these questions were presented to participants.

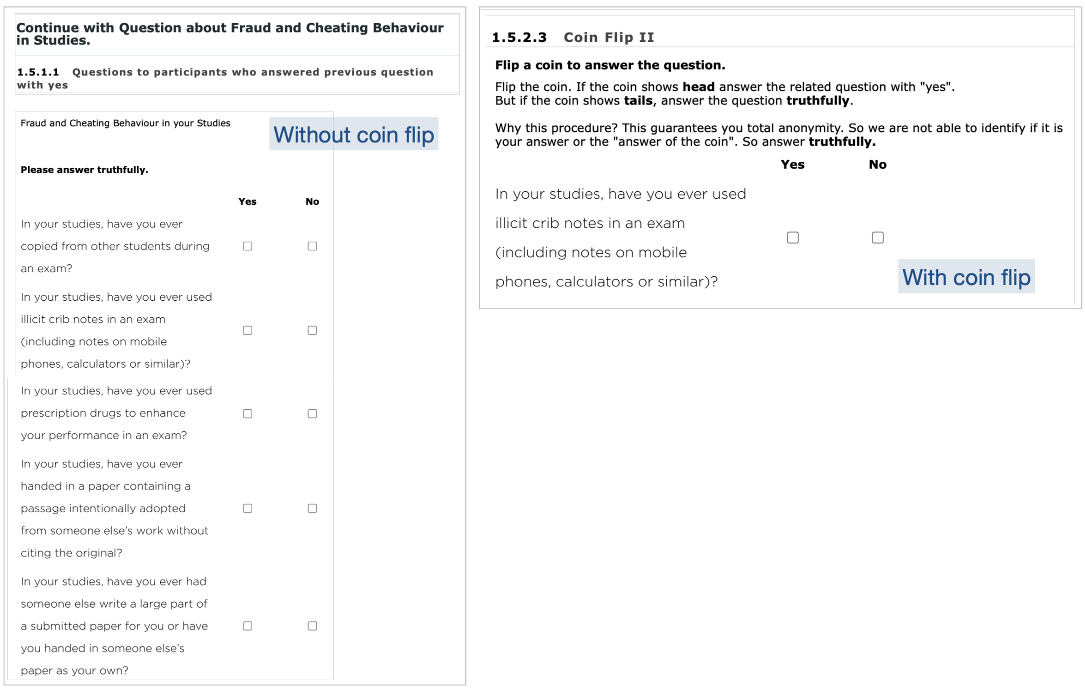


Figure 4.16: Dishonesty Questionnaire

In addition, the D-factor questionnaire, more precisely the 35-question version (D35), is documented in table 4.14.

Dataset	Split	Model	Model Architecture (Hyperparameters) - MLP-Classifiers
Repeated	Game	T	activation: 'tanh', alpha: 0.25, hidden_layer_sizes: (100, 100, 100), learning_rate: 'constant', learning_rate_init: 0.1, solver: 'adam', warm_start: True
		Y	activation: 'logistic', alpha: 0.1, hidden_layer_sizes: (150, 150, 150), learning_rate: 'constant', learning_rate_init: 0.2, solver: 'adam', warm_start: True
Repeated	P1 (Inspector)	T	activation: 'logistic', alpha: 0.001, hidden_layer_sizes: (150, 150, 150), learning_rate: 'constant', learning_rate_init: 0.2, solver: 'adam', warm_start: True
		Y	activation: 'tanh', alpha: 0.01, hidden_layer_sizes: (100, 100), learning_rate: 'constant', learning_rate_init: 0.2, solver: 'adam', warm_start: True
Repeated	P2 (Student)	T	activation: 'relu', alpha: 0.1, hidden_layer_sizes: (100, 100, 100), learning_rate: 'adaptive', learning_rate_init: 0.01, solver: 'sgd', warm_start: True
		Y	activation: 'relu', alpha: 0.0001, hidden_layer_sizes: (100, 100, 100), learning_rate: 'constant', learning_rate_init: 0.001, solver: 'sgd', warm_start: True
One-Shot	Game	T	activation: 'relu', alpha: 0.25, hidden_layer_sizes: (50, 50), learning_rate: 'constant', learning_rate_init: 0.2, solver: 'adam', warm_start: False
		Y	bootstrap: True, criterion: 'entropy', max_depth: 1, max_features: 'sqrt', min_samples_leaf: 30, n_estimators: 800
One-Shot	P1 (Inspector)	T	activation: 'tanh', alpha: 0.01, hidden_layer_sizes: (100,), learning_rate: 'invscaling', learning_rate_init: 0.2, solver: 'sgd', warm_start: False
		Y	activation: 'relu', alpha: 0.01, hidden_layer_sizes: (100,), learning_rate: 'constant', learning_rate_init: 0.001, solver: 'adam', warm_start: False
One-Shot	P2 (Student)	T	activation: 'tanh', alpha: 0.01, hidden_layer_sizes: (50,), learning_rate: 'constant', learning_rate_init: 0.1, solver: 'adam', warm_start: False
		Y	activation: 'logistic', alpha: 0.5, hidden_layer_sizes: (50, 50), learning_rate: 'constant', learning_rate_init: 0.1, solver: 'adam', warm_start: False

Table 4.12: First Stage Model Architecture Settings for the Selected Classification Models

Dataset	Split	Model Architecture (Hyperparameters) - Random Forest
Repeated	Game	n_estimators: 1200, max_depth: 5, max_features: 'sqrt', min_samples_leaf: 40, bootstrap: True, criterion: 'entropy', class_weight: 'balanced_subsample'
	P1 (Inspector)	n_estimators: 500, max_depth: 5, max_features: None, min_samples_leaf: 10, bootstrap: True, criterion: 'gini', class_weight: 'balanced'
	P2 (Student)	n_estimators: 700, max_depth: 5, max_features: None, min_samples_leaf: 60, bootstrap: True, criterion: 'gini', class_weight: 'balanced'
One-Shot	Game	n_estimators: 900, max_depth: 5, max_features: 'log2', min_samples_leaf: 55, bootstrap: False, criterion: 'entropy', class_weight: 'balanced_subsample'
	P1 (Inspector)	n_estimators: 1400, max_depth: 5, max_features: None, min_samples_leaf: 60, bootstrap: False, criterion: 'gini', class_weight: 'balanced_subsample'
	P2 (Student)	n_estimators: 500, max_depth: 5, max_features: 'log2', min_samples_leaf: 15, bootstrap: True, criterion: 'entropy', class_weight: 'balanced_subsample'

Table 4.13: Model Architecture Settings for the Random Forest Models

D35	Personality Trait	Reversed Score	Statements
1	Amoralism-crudelia	rev	It is hard for me to see someone suffering.
2	Psychopathy	reg	Payback needs to be quick and nasty.
3	Egoism	rev	All in all, it is better to be humble and honest than important and dishonest.
4	Spitefulness	reg	If I had the opportunity, then I would gladly pay a small sum of money to see a classmate who I do not like fail his or her final exam.
5	Machiavelianism	rev	Most people are basically good and kind.
6	Amoralism-crudelia	reg	My own pleasure is all that matters.
7	Psychopathy	reg	I'll say anything to get what I want.
8	Sadism	rev	Hurting people would make me very uncomfortable.
9	Egoism	reg	Never tell anyone the real reason you did something unless it is useful to do so.
10	Sadism	rev	If I ever tormented others, I would feel strong remorse
11	Machiavelianism	reg	I believe that lying is necessary to maintain a competitive advantage over others.
12	Self-centeredness	rev	I feel sorry if things I do upset people.
13	Amoralism-frustralia	reg	A person should use any and all means that are to his advantage, taking care of course, that others do not find out.
14	Psychopathy	reg	People who mess with me always regret it.
15	Narcissism	rev	In principle, everyone is worth the same.
16	Sadism	rev	I cannot imagine how being mean to others could ever be exciting.
17	Egoism	reg	To make money there are no right and wrong ways anymore. Only easy and hard ways.
18	Psychopathy	rev	I don't want people to be afraid of me or my impulses.
19	Psychological Entitlement	rev	I do not deserve more things in life than others.
20	Amoralism-frustralia	reg	I would like to make some people suffer, even if it meant that I would go to hell with them.
21	Machiavelianism	reg	It's wise to keep track of information that you can use against people later.
22	Self-centeredness	reg	I'm not very sympathetic to other people or their problems.
23	Narcissism	rev	It does not give me much pleasure to see my rivals fail.
24	Psychopathy	rev	I make a point of trying not to hurt others in pursuit of my goals.
25	Moral Disenrangement	reg	People who get mistreated have usually done something to bring it on themselves.
26	Amoralism-crudelia	reg	Why should I care about other people, when no one cares about me?
27	Sadism	rev	I avoid humiliating others.
28	Machiavelianism	rev	Most people deserve respect.
29	Psychological Entitlement	reg	Someone who hurts me cannot count on my sympathy.
30	Spitefulness	reg	I would be willing to take a punch if it meant that someone I did not like would receive two punches.
31	Greed	rev	For most things, there is a point of having enough.
32	Psychopathy	reg	Success is based on survival of the fittest; I am not concerned about the losers.
33	Narcissism	rev	I do not mind sharing the stage.
34	Amoralism-crudelia	reg	Doing good deeds serves no purpose; it only makes people poor and lazy.
35	Sadism	rev	Making people feel bad about themselves does not make me feel any better.

Table 4.14: D-Factor Questionnaire (D35)

5. Trust in the Machine: How Contextual Factors and Personality Traits Shape Algorithm Aversion and Collaboration

Authors

Vinícius Ferraz, Leon Houf, Thomas Pitz, Christiane Schwieren & Jörn Sickmann

Abstract

This paper investigates the interplay between contextual factors, personal variables, and algorithm aversion in decision delegation behavior. In an experimental setting with four treatments —baseline, explanation, payment, and automation— subjects chose whether to delegate decisions to an algorithm, considering hidden expected values. Employing Random Forests, Gradient Boosting Machines, and causal analysis with the Uplift Random Forest, we probed key algorithm aversion drivers. In the personal dimension, we assessed Big Five Personality Traits, Locus of Control, Generalized Trust, and demographics. We find that payment reduced delegation, while full automation promoted it. Factors like Age, Extraversion, Openness, Neuroticism, and Locus of Control consistently emerged as significant in shaping delegation decisions. Female participants demonstrated a stronger reaction to algorithmic mistakes. This study offers insights for crafting user-centric AI design to enhance cooperation and minimize aversion.

Keywords

Algorithm Aversion, Human-Computer Interaction, Decision Behavior, Machine Learning, Causal Inference

5.1 Introduction

Driven by technological advancements, data availability, and computing power, intelligent systems powered by Artificial Intelligence (AI) have become common in our society, largely due to their transformative potential (S. J. Russell, 2010). AI simulates human behaviors like learning and decision-making (McCarthy, 2007). AI's ability to efficiently process vast amounts of data, inform decisions, and automate processes has led to its widespread adoption (Azucar, Marengo, & Settanni, 2018).

However, technological shifts can lead to new social phenomena like algorithm aversion, characterized by the reluctance to use algorithms in decision-making, despite their superior ability to undertake certain tasks (Dietvorst, Simmons, & Massey, 2015; Ku, 2020). The extensive body of literature emerging in a relatively short period reveals an intricate mechanism with various factors that can influence aversion or appreciation of algorithms, demonstrating the complexity of achieving a common understanding of the underlying reasons for this behavior.

As a consensus in the literature, context and personal elements significantly shape an individual's willingness or aversion to delegate decisions to an automated system. Building on this concept, we explore these two impact dimensions in an experimental study, applying a simplified multi-armed bandit problem. In the experiment, subjects repeatedly choose from three options with hidden expected values, aiming to identify the superior option. At each period, they can delegate decisions to a Reinforcement Learning algorithm. For a holistic understanding of this behavior, this study delves into the environmental dimension by investigating the impact of explainability, costs, and full task automation. Concurrently, we assess the personal dimension by examining personality traits commonly associated with algorithm aversion, such as the big five, locus of control, generalized trust, and demographic information. Despite growing awareness of algorithm aversion, there remains a need for more extensive research; therefore, we focused on these psychological and contextual measures inspired by suggestions and recommendations for further research in (Burton, Stein, & Jensen, 2020; Mahmud, Islam, Ahmed, & Smolander, 2022). This paper aims to contribute to understanding how to design systems that enable fruitful interactions between humans and computers.

Experimental evidence on algorithm aversion and appreciation varies significantly across domains and contexts. Studies have found differing levels of human interaction with automated agents based on factors such as task context, performance expectations, and agent roles (Chugunova & Sele, 2022). Studies in financial and investment contexts highlight reluctance to fully surrender decision-making authority to automated agents despite their superior performance (Filiz, Judek, Lorenz, & Spiwoks, 2022; Gaudeul & Giannetti, 2023; Logg, 2017). The presence of human errors and significant decision outcomes seem to exacerbate algorithm aversion (Dietvorst et al., 2015; Filiz, Judek, Lorenz, & Spiwoks, 2021). Yet, showcasing an AI-based system's learning ability (Berger, Adam, Rühr, & Benlian, 2021) or exerting time pressure can mitigate this aversion (Jung & Seiter, 2021). Notably, the moral implications of decisions also play a role. In morally charged decisions, people often prefer the discretionary scope of human decision-makers (Jauernig, Uhl, & Walkowitz, 2022), and in situations where discrimination is possible, people prefer algorithmic evaluation (Jago & Laurin, 2022). However, there are instances of preference for algorithmic over human advisory, influenced by factors such as how the expertise of the algorithm is framed against a human (Candrian & Scherer, 2022; Hou & Jung, 2021; Logg, Minson, & Moore, 2019). In summary, experimental studies on human-machine collaboration and algorithm aversion point to the complexity of these phenomena, influenced by a range of factors from decision consequences and task complexity to decision context framing and perceived algorithm expertise. For comprehensive and interdisciplinary literature collections on algorithm aversion, systematic reviews are provided in Jussupow, Benbasat, and Heinzl (2020), Burton et al. (2020), and Mahmud et al. (2022).

Given the complexity of the phenomenon, we began our methodological approach with statistical and regression analysis to understand treatment differences and explore variable relationships. We then used machine learning and causal inference techniques, including Logistic Regressions, Random Forest, Gradient Boosting Machines, and Uplift Random Forest classifiers, to probe the nuanced nature of decision delegation behavior. Contextual factors like payment and automation notably affected delegation, with payment reducing and full automation boosting its likelihood. Key personal factors influencing delegation across models were age, extraversion, openness, neuroticism, and locus of control. This paper documents the intricate relationships between individual traits, contextual conditions, and delegation behavior, providing a nuanced understanding of algorithm aversion within the boundaries of an experimental construct.

5.2 Experimental Design

The experimental setting employed a between-subject design, utilizing a simplified version of the multi-armed bandit problem (Robbins, 1952). Our design finds parallels in previous works, notably by Hoelzemann and Klein (2021), who also examined human interactions with bandit-based decision-making scenarios. The primary task involved participants repeatedly choosing one of three options labeled as "products" over 40 periods. The experiment was conducted online, where participants were instructed to select from three products, each with distinct hidden quality levels that represented their expected values, translated into the probability of receiving a payoff from the chosen option. The three variants of quality were low (50% chance of payoff), medium (70% chance of payoff), and high (90% chance of payoff). These probabilities were randomly assigned to products 1 to 3 at each participant's onset and remained constant throughout the experiment. Through repeated choices, the expected goal was for the participants to identify the high-quality product that would maximize their total payoffs. After each selection, participants received feedback on the outcome of their decision. In each round, participants had the option to delegate the decision to an algorithm. After reading the instructions, we asked participants about their perception of using algorithms for decision-making in regular tasks. The responses were categorized as positive, neutral, or negative. This response was used as a variable in the study, referred to as *perception*.

The basic framework described above is established as the "baseline" treatment. We further introduce three treatments with different contexts — Explanation, Payment, and Automation — to investigate the impact of explainability and transparency, willingness to pay, and complete task automation on delegation behavior. We aim to better understand user preferences and friction points in algorithmic decision-making by examining these factors. In all treatments, we employ an attention check in a given round by displaying an animal picture below the task, which participants had to identify by the end of the task. Information about the design and the actual experiment screens are documented in appendix 5.6.4.

5.2.1 Explanation Treatment

As discussed in numerous studies, transparency and explainability are key factors affecting the acceptance of algorithmic decision support. Algorithm complexity often presents these tools as "black boxes," undermining their acceptance due to the lack of understanding (De Bruyn, Viswanathan, Beh, Brock, & von Wangenheim, 2020; Enholm, Papagiannidis, Mikalef, & Krogstie, 2021; Miller, 2019; Trocin, Mikalef, Papamitsiou, & Conboy, 2021; Vlačić, Corbo, e Silva, & Dabić, 2021; Y. Zhang, Chen, et al., 2020).

The inherent complexity in high-performing computational models poses a dilemma between accuracy and transparency, as the intricacy of these models could challenge the public's comprehension (Gilpin et al., 2018; Gunning, 2017; Herm, Heinrich, Wanner, & Janiesch, 2022). This complexity underscores the ongoing challenge practitioners face in maintaining explainability (Castelluccia & Le Métayer, 2019), necessitating accessible explanations irrespective of the chosen approach. Institutions and regulators also emphasize the need for transparent algorithmic decisions (Goodman & Flaxman, 2017).

We tested the information-sharing impact on delegation in this *explanation* treatment, in which participants had access to a description of the algorithm used in the product selection task. The description was supposed to be non-technical and to transmit the essence of the method behind reinforcement learning to the subjects. In the primary experiment page, the following text is displayed in a text box with a prominent design: *"Reinforcement Learning: the algorithm calculates probabilities and chooses an alternative based on the success of choices in previous rounds"*. The description text remained visible during the experiment.

5.2.2 Payment Treatment

Exploring the less examined aspect of financial incentives in algorithm aversion, people might hesitate to pay for transparent AI if costs surpass perceived benefits (König, Wurster, & Siewert, 2022). During crises, the appeal for robo-advisors—and hence the willingness to pay—escalates due to the need for financial advice (Ben-David & Sade, 2001). Similarly, radiologists are ready to pay for AI tools that expedite diagnostics (von Wedel & Hagist, 2022).

We investigate payment's role in algorithm aversion by assigning a payment requirement to algorithmic support, termed *payment* treatment. Here, participants were informed that while they can delegate decisions to an algorithm, each delegation carries a cost of 0.10 points (one-tenth of a point), aiming to introduce the psychological aspect of payment in a way that participants easily understand. The goal was to simply introduce payment as a contextual variable to gauge its impact, not to explore the complexities of differential willingness to pay. The cost incurred for a decision effectively restricts algorithm support to a pay-per-use basis. The points deduction reduces the expected values of the products by the same amount, introducing a "loss" for rounds where payoffs do not materialize, as the amount is subtracted from the participant's total points.

5.2.3 Automation Treatment

The task complexity may induce people towards higher acceptance of algorithmic decisions (Bogert, Schecter, & Watson, 2021). Bucklin, Lehmann, and Little (1998) argue that from a human standpoint, full, compared to partial, automation of decision-making processes can be very desirable in terms of efficiency, such as improving productivity, and effectiveness, for better resource allocation. In essence, the action of delegating the decision is already a form of automation, as the algorithm calculates and selects the best option based on past data. We advance this process by further automating it, thereby reducing the overall task burden. In this way, one can analyze the subjects' behavior toward the delegation of discrete decisions compared to the delegation of the complete task.

In the *automation* treatment, the algorithm takes over the repetitive task of product selection for 40 periods, easing the participants' effort. Unlike previous treatments requiring round-by-round delegation decisions, this feature allows continuous selection without active involvement. Participants could toggle automation on or off at any stage. If they opted for delegation, they had a 5-second window to override the decision, redirecting them to the primary selection interface. Feedback remained available post each round.

5.2.4 Personal Dimension

Algorithm aversion can be significantly impacted by personal factors such as psychological aspects, personality traits, demographic features, and algorithm/task familiarity (Mahmud et al., 2022). For instance, individuals with an internal locus of control tend to resist human and AI suggestions (Sharan & Romano, 2020), and neuroticism correlates with lower trust ratings. Delegation to algorithms increases when information scarcity is present and among extroverted individuals (Goldbach, Kayar, Pitz, & Sickmann, 2019). Trust in algorithms is not static but can evolve with personal experiences (Fenneman, Sickmann, Pitz, & Sanfey, 2021), which similarly impacts attitudes toward autonomous transport (Goldbach, Sickmann, Pitz, & Zimasa, 2022).

Broadening our research to encompass both contextual and personal aspects of algorithm aversion, we incorporate demographic data, the Big Five Personality traits, Locus of Control, and trust levels into our analysis. The Big Five Personality Traits offer an encompassing view of human personality (L. R. Goldberg, 1990), while Locus of Control illustrates an individual's belief in their power over life events (Rotter, 1966). Generalized trust signifies an individual's confidence in the reliability and benevolence of others (Yamagishi & Yamagishi, 1994). After completing the selection task, participants proceeded to this series of personality questionnaires, including control questions (see appendix 5.6.4).

5.3 The Algorithm: Reinforcement Learning Implementation Framework

The term "algorithm" has various definitions across different fields. Computer science typically defines it as a step-by-step procedure or set of rules used to perform tasks (Cormen, Leiserson, Rivest, & Stein, 2001). In the context of algorithm aversion, it often refers to decision-making tools that assist humans in making choices or predictions (Dietvorst et al., 2015).

A variety of algorithms could be applied to the task of repeatedly selecting alternatives that maximize one's payoffs. In our design, we aimed to allow participants to observe the algorithm's training and improvement process throughout the task while keeping it simple enough for participants in the explanation treatment to understand its core mechanism in just a sentence or two. As a result, we chose the Reinforcement Learning (RL) model, a class of solution methods well-suited for learning-based and sequential problems.

Reinforcement learning is typically framed as an optimization problem, with the goal of identifying optimal actions based on defined criteria (Barto, 1997). The model's framework is designed to map situations to actions in a way that maximizes rewards, as defined by Sutton and Barto (2018). Key components of reinforcement-based models include a set of choices or actions, a mechanism for receiving feedback associated with each choice, an updating rule that adjusts previous beliefs or estimates of each choice's expected value based on the feedback, and a decision rule that determines the probability of selecting each choice based on current beliefs. Our model is based on Erev and Roth (1998) 's implementation, which incorporates the concept of attractions, or weights attached to strategies that represent the perceived value associated with specific choices (C. Camerer & Hua Ho, 1999). Our implementation assigns an attraction value to each product, which is updated after a decision is made using a learning rule. The attractions are transformed into probabilities of choice using a softmax function. A formalization of the algorithm is presented in appendix 5.6.1.

The embedding of this algorithm in the experiment generates one instance of reinforcement learning for each participant, which starts with no pre-training or bias. The attraction values are initialized at 0, and the algorithm learns from participant choices and its own choices over time, making the learning process for humans and algorithms comparable.

5.4 Results

In this section, we conduct a comprehensive six-stage analysis of decision delegation to an algorithm, exploring its contextual, behavioral, and personal dimensions. We begin with an overview of our sample information and attention analysis, followed by an examination of delegation behavior across different treatments. We then use regression methods to identify significant predictors of delegation behavior and machine learning methods for a nuanced understanding of algorithm aversion. We incorporate causal inference methods to clarify causal relationships, analyze participants' reactions to algorithmic failures, and measure the algorithm's performance under varying conditions. This multifaceted approach provides a detailed understanding of the complex phenomenon of algorithm aversion[1].

5.4.1 Sample Information and Attention Analysis

A total of 358 participants took part in our online experiment. Subjects were evenly distributed across the four treatments, with approximately 89 to 91 participants per treatment. On average, the experiment took 11 minutes to complete, and participants earned

[1]This research project was pre-registered in AsPredicted.org, with the ID 119401.

between 4 and 10 euros, with an average of 6.13 euros. Demographically, the sample was 52.7% female. Participants were primarily from Germany (51%), with the remaining individuals representing various nationalities. Most participants (73.2%) were from the Rhine-Waal University of Applied Sciences, while 26.8% were from Heidelberg University (both in Germany), aged between 18 and 47 years old; the mean age was 25. Among the subjects, 19% were economics students; the rest were from various other academic disciplines, of which 21% came from STEM majors. The self-reported perception values were 46.6% positive, 43.9% neutral, and 9.5% negative.

We analyzed participants' attention, particularly focusing on the automation treatment, to determine if active supervision of the algorithms' decisions persisted in a fully automated task. To measure this, we calculated the total time the web page was active in the subjects' browsers. Additionally, we implemented attention-check questions in both the experimental task and the personality questionnaires. The results are summarized in the table 5.4.1; these values do not account for the first round, which includes the time of reading the instructions.

Treatment	Average Active Time (s)	Animal Question (frequency correct)
Baseline	9.6	0.88
Explanation	10.3	0.89
Payment	9.0	0.85
Automation	11.2	0.55

Table 5.1: Attention metrics for all treatments

The active time analysis showed consistent results across all treatments, with participants spending an average of 9 to 11 seconds per round. A second attention check involved identifying an animal that appeared during the final rounds, revealing decreased attention in the automated treatment. Even though the screen was active, fewer people in the automated treatment seemed to monitor the task closely. We included an attention self-report question in the automated treatment, especially asking if the subject had supervised the algorithm's decisions during the task. 76% of them answered yes, which deviates from the 55% of participants that got the animal question correct. 15% answered no, and 9% answered not applicable. The delta suggests an overreporting of the attention and supervision levels in the automated treatment. Four control questions were embedded in the personality tests, with 78% of participants answering all four correctly and 93% answering at least three correctly, indicating attentive reading.

5.4.2 Delegation Behavior and Treatment Effects

We measured the frequency of delegating decisions to the algorithm in each treatment. The absolute frequency of delegation in each treatment is documented in table 5.4.2.

In the baseline treatment, we observed a balanced split, where about half of the decisions were delegated across participants and rounds. The information shared in the explanation treatment only slightly increased the number of delegation decisions. The introduction of

Treatment	Frequency of Delegation
Baseline	53.02%
Explanation	58.37%
Payment	27.87%
Automation	66.07%

Table 5.2: Absolute frequencies of delegation across the four treatments

payment sharply decreases, and the possibility for automation increases the willingness to allow the algorithm to decide.

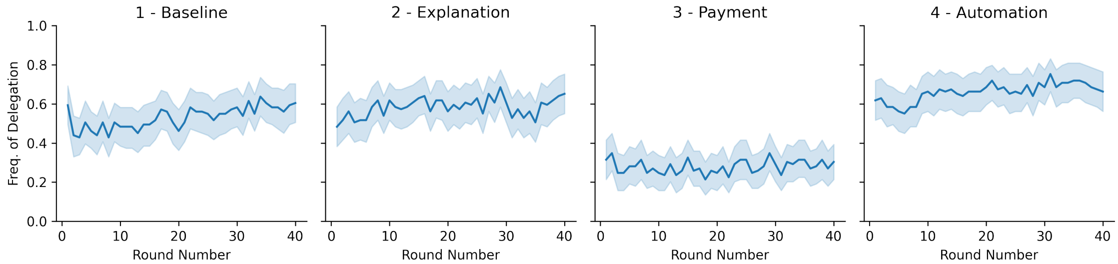


Figure 5.1: Mean Frequencies of Delegation Over Time

Figure 5.1 displays the overall delegation frequencies over time, where the distributions are consistent across treatments and relatively constant, without any large variations in the decision behavior between rounds. We aggregated the experimental data on a participant level to test these findings for statistical significance. Each participant's cumulative delegation frequency over 40 periods is treated as an independent observation. The distributions of these relative frequencies of delegation are displayed in the histogram in figure 5.2.

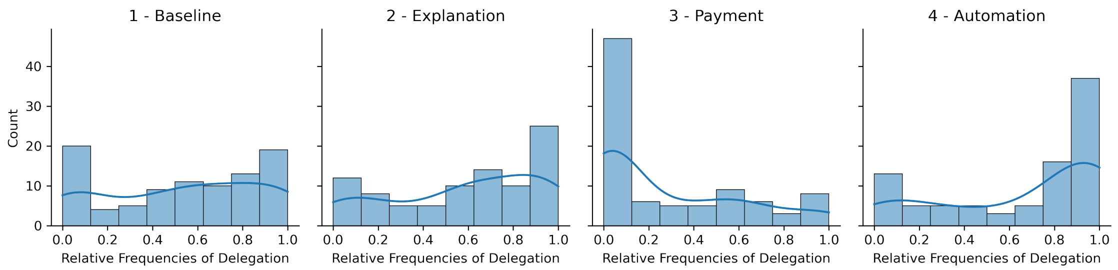


Figure 5.2: Histogram of Participants Cumulative Delegation Frequencies

As anticipated, the highest delegation frequencies occur in automation and the lowest in payment treatments. The baseline and explanation treatments exhibit a more even distribution of subjects' delegation behavior. We employed a Kruskal-Wallis test (Kruskal & Wallis, 1952), a non-parametric statistical test comparing the medians of several independent samples. With a test statistic of 52.67 and a p-value < 0.001, the results indicate a significant difference between the medians of the four independent treatment samples.

While the Kruskal-Wallis test reveals significant differences, it does not provide detailed insights into these differences between the samples. Consequently, we employed a Dunn posthoc test (Dunn, 1961) to identify significant pairwise differences between samples. The p-values for these comparisons are in table 5.4.2.

	Baseline	Explanation	Payment	Automation
Baseline	1	0.373	<0.001	0.009
Explanation	0.373	1	<0.001	0.090
Payment	<0.001	<0.001	1	<0.001
Automation	0.009	0.090	<0.001	1

Table 5.3: Dunn posthoc test results, p-values for pairwise treatment comparisons

In summary, these results suggest significant differences between the medians of baseline, payment, and automation, as well as between explanation and payment. There is no significant difference between the medians of baseline and explanation or between explanation and automation. The payment feature was the most influential regarding the willingness to delegate.

The contextual findings highlight the influence of different treatment conditions on the delegation behavior of participants. The baseline and explanation treatments led to a more even distribution of delegation behavior. On the other hand, the payment treatment had a considerable negative impact on the willingness to delegate. The automation treatment led to the highest frequency of delegation among the four treatments, demonstrating the importance of reducing the involved workload in a task in encouraging algorithm-based decision-making. Overall, these results underscore the significance of understanding and addressing the factors that affect delegation behavior to design more effective human-algorithm collaborations and decision-making processes.

5.4.3 Incorporating the Personal Dimension - Regression Analysis

The design of our treatments provides insights into how exogenous factors influence delegation behavior. However, individual factors also play a significant role in algorithm aversion, as widely discussed in the literature. In this section, we examine the binary action of delegating a decision in relation to treatment conditions and personal factors, including personality test scores, gender, education, and self-reported perception (as explained in chapter 5.2). Categorical values were encoded as binary dummy variables.

Although correlations between the variables under investigation and delegation are primarily weak, they are highly significant (full correlation results are reported in 5.6.2, appendix 5.6.2). To further explore and quantify these relationships, we constructed a logistic regression model including demographic and personal information as independent variables. The model results are summarized in table 5.4. A critical remark in the regression modeling is that we use the entire experiment's dataset: every decision from each participant at each round. Due to repeated choices made by the same individuals across 40 periods, we clustered the standard errors on the participant level. This approach accounts for intra-participant correlation, considering potential influences from unobserved individual factors or shared experiences, as per Bertrand et al. (2004) 's reasoning.

Variable	Coefficient	Standard Error	p-value
Constant	-0.525	1.013	0.605
Explanation	0.252	0.207	0.223
Payment	-1.012	0.235	*** < 0.001
Automation	0.515	0.234	*0.027
Female	-0.144	0.179	0.421
Age	-0.009	0.018	0.596
STEM	0.267	0.227	0.238
Business & Economics	-0.181	0.201	0.37
Extraversion	0.04	0.059	0.497
Agreeableness	0.036	0.073	0.627
Conscientiousness	0.137	0.085	0.106
Neuroticism	-0.047	0.072	0.513
Openness	-0.008	0.087	0.926
Internal LoC	0.057	0.102	0.578
External LoC	0.054	0.106	0.614
Generalized Trust	0.067	0.066	0.307
Perception	-0.368	0.14	**0.009

Note: Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 5.4: Logistic Regression Results - Delegation

The logistic regression model provides several insights into the effects of treatments and personality traits on delegation and also reinforces the findings in chapter 5.4.2. Initially, the automation treatment exhibits a positive and statistically significant impact on delegation ($p = 0.027$), suggesting that automating tasks encourages individuals to delegate. Conversely, the payment treatment displays a negative and statistically significant influence ($p < 0.001$), implying that requiring payment could discourage delegation. The explanation treatment, although positive, is not statistically significant ($p = 0.223$). Regarding personal variables, the only statistically significant effects are observed for perception ($p = 0.009$), which negatively impacts delegation, suggesting that an increase in negative perception about algorithms is correlated with a lower likelihood of delegation. Other variables, including gender, age, field of study, and personality traits, do not exhibit statistically significant effects on delegation in this model. A second regression model, including interaction terms, is reported in appendix 5.6.2, in which payment loses its significance, and Internal Locus of Control becomes significant. Quantile regression models applied to cumulative delegation frequencies (shown in figure 5.2) showed similar significance and coefficients to logistic regression, despite a marginally better fit. See appendix 5.6.2 for full details.

In conclusion, examining personality traits and algorithm aversion uncovers the influence of individual factors and treatment conditions on delegation behavior. A critical insight from this analysis is the existence of intricate relationships between various traits. Interaction terms offer a more comprehensive understanding of the relationships between variables and delegation behavior by accounting for the dependence of some variables' effects on the values of other variables. Gaining insights into these relationships can aid in comprehending how diverse behavioral profiles respond to algorithmic systems.

5.4.4 Machine Learning for Delegation Behavior Analysis and Causal Inference

To understand whether the personal and contextual pieces of information are helpful in predicting the delegation behavior in such a case, we tested a few prediction techniques using the same variables scheme, that is, predicting the binary outcome of the delegation decision possibility using the treatments, personality, and demographic data.

C. F. Camerer (2018) highlights the benefits of applying machine learning to model behavior, emphasizing its potential for improved predictive accuracy, handling large datasets, capturing non-linear relationships, and adaptability. Additionally, machine learning enables personalization and fosters cross-disciplinary insights, contributing to a better understanding of human decision-making and facilitating more effective interventions across various domains.

The logistic regression model, as detailed in chapter 5.4.3, offers limited insights into the complex interplay of our variables, accounting for only about 9% (pseudo R-squared) of the variation in delegation decisions. Given the absence of clear linear relationships and the complexity of the data, we turn to more sophisticated methods. We employ machine-learning models to examine the overall impact of variables on predicting delegation, followed by causal machine learning models to separate treatment effects from the personal covariates. In the subsequent models, we refer to within-sample predictions, using 80% of the sample for model training and the other 20% to generate and test predictions. Methodological formalizations for the adopted methods can be found in appendix 5.6.1, and technical model implementation remarks in appendix 5.6.3.

5.4.5 Predicting Delegation Behavior

If we use our logistic regression coefficients to generate predictions, the model yields an accuracy score of 0.62, meaning 62% of the delegation decisions were classified correctly, not far from a random baseline. This relatively low accuracy might be due to several factors influencing the results that have yet to be accounted for or the failure of the model to capture complex relationships between the variables. To deepen the understanding of these variables' relationships and the possibility of generating predictions for algorithm aversion behavior using contextual and personal information, we resort to the machine learning techniques Random Forest and Gradient Boosting Machines.

Research shows successful predictions of behavioral elements using personality traits, characteristics, and environmental data. Balakrishnan, Khan, Fernandez, and Arabnia (2019) used psychometric test data, including Big 5 and Dark Triad, and Twitter features to predict cyberbullying accurately. Guntuku, Yaden, Kern, Ungar, and Eichstaedt (2017) employed machine learning to predict mental health status based on social media and personality data. Similarly, Stachl et al. (2017) used personality traits to predict smartphone usage behavior. Saltık, Söyü, Değirmen, Şengönül, et al. (2023) combined reaction time, psychological attributes, and personality traits to predict Loss Aversion Bias, supporting Kahneman's "Thinking Fast and Slow" theory (Kahneman, 2011). These studies demonstrate the potential of machine learning models in similar prediction tasks.

Breiman (2001) introduced the Random Forest model, an ensemble learning method designed for classification and regression problems. The algorithm works by creating multiple decision trees, each of which 'votes' on an answer. In a classification problem such as ours, the Random Forest chooses the class that gets the most votes from all the trees. The key idea behind Random Forest is to create a "forest" of diverse decision trees constructed from random subsets of training data and features. This approach helps increase the model's robustness, reduce overfitting, and improve overall predictive accuracy. The Random Forest algorithm is particularly useful for binary classification problems because it can handle non-linear relationships between the input features and the output variable. It can also handle missing values and outliers in the input data and estimate the importance of each input feature in the prediction (Liaw, Wiener, et al., 2002).

In a similar manner, Gradient Boosting Machines (GBMs) are a class of ensemble learning algorithms that build a robust model by iteratively adding weak learners, typically decision trees, to minimize a loss function. The algorithm focuses on correcting the errors of the previous tree by training on the residuals, effectively improving the overall model's performance, as defined in J. H. Friedman (2001).

As per definitions in Breiman (2001) and J. H. Friedman (2001), Random Forest and GBMs are ensemble learning methods for similar purposes. The main difference lies in their approach to building the ensemble of decision trees. Random Forest constructs multiple trees independently and in parallel, combining their predictions through averaging or majority voting. It uses bagging (Bootstrap Aggregating) to create diverse trees by resampling the dataset with replacement. In contrast, GBM constructs trees sequentially, with each new tree trying to correct the errors made by the previous tree. It utilizes a technique called boosting, where trees are combined through a weighted majority vote, and the weights are determined by minimizing a loss function during the training process. We apply both methods for comparable results but with distinct processes, enabling comparing and validating the findings from the generated predictions to assess our findings' consistency. In each model, feature importances highlight the significance of each feature in predicting the target variable. Figure 5.3 presents an overview of the feature importances.

Both models have been cross-validated during parameter fitting and training using the KFold method to avoid overfitting (details in appendix 5.6.3). In this process, we split the training data into a number of subsets or "folds." We train the model on the remaining data for each fold and test it on this fold. This process is repeated for each fold, allowing us to assess the model's performance based on its ability to predict new data (Berrar, 2019; Kohavi et al., 1995). Furthermore, with an equivalent objective as clustering the regression errors on a participant level (chapter 5.4.3), we aggregated the participant observations here using the GroupKFold variant, which ensures instances from the same participant either in the training set or the test set. This approach safeguards against data leakage and maintains a realistic estimate of the model's performance, especially when observations within the same group (in this case, participant) are correlated.

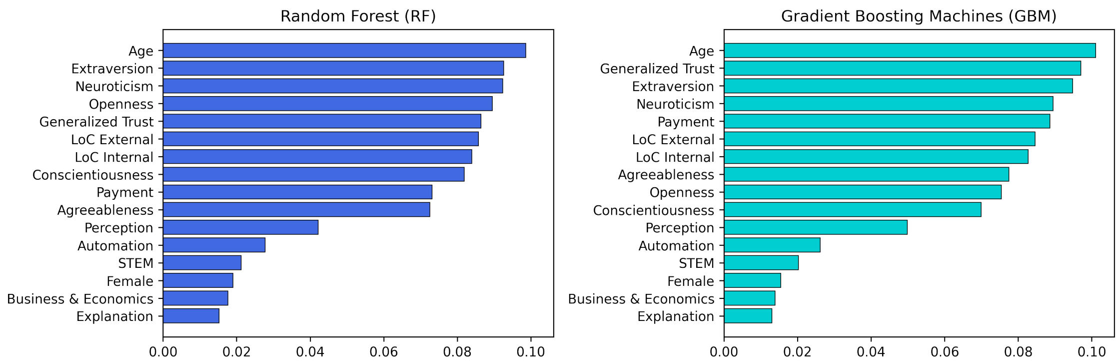


Figure 5.3: Machine Learning Models Feature Importances

According to the Random Forest and Gradient Boosting Machine models, the decision to delegate to algorithms is influenced by a complex mix of individual characteristics and contextual factors. Age consistently emerges as the most significant variable in both models, reflecting its significant role in shaping comfort with algorithmic delegation. Similarly, Neuroticism and Extraversion — two Big Five personality traits — feature prominently, signifying their impact on delegation tendencies.

Apart from these, the Locus of Control, both internal and external, appears to influence delegation decisions, although they are more pronounced in the Random Forest model. Contextual factors, like payment and automation, also emerge as crucial determinants across both models. Intriguingly, automation is more influential in the GBM model, suggesting a more substantial bias towards delegation in fully automated scenarios. Gender, education, and the Explanation context appear to have minimal impact in both models.

These findings underscore the intricate dynamics governing decision delegation, with no single factor having a dominating influence. Instead, a nuanced interplay of various individual and contextual elements appears to guide the decision to delegate to algorithms.

We evaluated the Logistic Regression (LR), Random Forest (RF), and Gradient Boosting Machine (GBM) models using four metrics: Accuracy, Precision, Recall, and F1 score. Accuracy calculates the proportion of correctly classified instances. Precision quantifies how well the model correctly identifies positive instances. Recall gauges the model's ability to detect positive instances among actual positives. The F1 score, a blend of precision and recall, is the harmonic mean of these two metrics (Powers, 2020; Sokolova & Lapalme, 2009). As summarized in Table 5.4.5, both RF and GBM outperformed LR in predictive power, with RF achieving slightly superior performance across all metrics. This outcome highlights the efficacy of tree-based models for our classification problem.

In addition, a Receiver Operating Characteristic (ROC) curve provides a graphical representation of a classifier's performance across varying decision thresholds (figure 5.4). The Area Under the ROC Curve (AUC-ROC) measures the overall performance of a binary

	LR	RF	GBM
Accuracy	0.6210	0.8332	0.8325
Precision	0.6112	0.8185	0.8120
Recall	0.7018	0.8730	0.8730
F1-score	0.6534	0.8414	0.8415

Table 5.5: Prediction Performance metrics

classifier. It ranges from 0 to 1, with higher values indicating better performance. A value of 0.5 indicates a random classifier (dashed line), and 1 indicates a perfect classifier. The ROC area quantifies how well the classifier can distinguish between the positive and negative classes, regardless of the choice of classification threshold (Bradley, 1997; Fawcett, 2006). In the overall analysis, and in line with previous performance metrics, the LR model is surpassed by the other models, with the RF model showing a slight edge. The high scores achieved by both the RF and GBM models affirm their ability to explain the data, enhancing the reliability of the interpretations documented in our study.

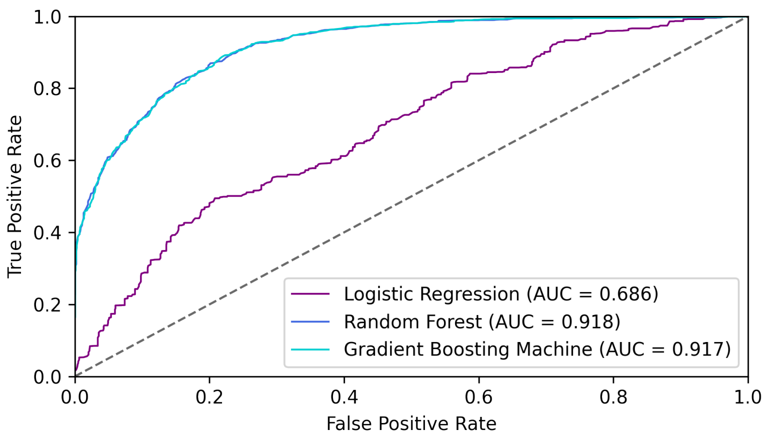


Figure 5.4: ROC Curves for All Models

Although logistic regression provided valuable insights into the direction and significance of individual variables, its ability to handle the complex data relationships in our study was limited. We explored machine learning techniques to capture these relationships better, specifically Random Forest and Gradient Boosting Machines. Both models significantly outperformed logistic regression regarding accuracy, precision, recall, and F1 score, with the Random Forest model having a slight edge in accuracy over the GBM. Both models consistently highlighted the same features, such as payment, extraversion, and neuroticism, as key influencers in delegation decisions.

5.4.6 Causal Inference and Heterogeneous Treatment Effects - Uplift Random Forest

To further understand the factors influencing decision delegation to algorithms, we now focus on disentangling the effects of the treatment conditions from personal data. While regression and machine learning models have provided insights, they combine all variables, not distinguishing between treatment conditions and personal characteristics effects. Hence, we use causal inference to uncover how treatment effects vary across different subgroups within our sample, focusing on estimating the expected change in the outcome as a result of the intervention. This approach allows us to measure heterogeneous treatment effects and identify the subset of individuals most influenced by the treatment conditions, given their characteristics. To this end, we resort to Uplift Modeling.

Uplift Modeling, a branch of causal inference, models the impact of incremental treatment effects on individuals' behavior (N. J. Radcliffe & Surry, 2011). Early applications of similar methods can be seen in N. Radcliffe and Surry (1999). For a comprehensive definition and literature review on machine learning problems and applications, see Gutierrez and Gérardy (2017); N. J. Radcliffe and Surry (2011).

We employ the Uplift Random Forest Algorithm, an ensemble learning method that uses the random forest algorithm to estimate the causal effect of a treatment or intervention on individual outcomes (Guelman, Guillén, & Pérez-Marín, 2012, 2015). The uplift random forest classifier (Sołtys, Jaroszewicz, & Rzepakowski, 2015) incorporates the treatment indicator as a covariate to capture differential effects and uses other covariates to estimate individual treatment effects. The model is tuned using the same cross-validation technique described in 5.4.5, with details in appendix 5.6.3.

Treatment effects can be evaluated at an individual level by computing uplift scores. These scores represent the predicted likelihood of delegation for each observation under each treatment scenario, essentially providing a probabilistic estimate of how a participant would behave if they were subjected to a specific treatment. The distributions of these predicted likelihoods are plotted in figure 5.5. The trend observed in this analysis follows the initial assessment of the treatment effects (chapter 5.4.2) in reference to the baseline. Payment negatively impacts the likelihood of delegation, whereas explanation has a slight positive effect, and automation has a more pronounced positive effect. Each treatment's computed average treatment effects are payment $= -0.26$, explanation $= 0.05$, and automation $= 0.12$.

Feature importance can also be extracted from this model, with a slightly different meaning. Unlike traditional classification models, in Uplift models, feature importance does not directly equate to the effect of a feature on the outcome but rather its influence on the treatment effects. In other words, an essential feature in the model translates to the influence on the change in the likelihood of delegation mediated by the treatment. These values are presented in figure 5.6.

In the Uplift Random Forest model, Age, Openness, External Locus of Control, Extraversion, and Agreeableness significantly influence the treatment effectiveness on delegation behavior. Other variables like Internal Locus of Control and Conscientiousness also play a role, but their influence is moderate. On the other hand, Gender, Perception, and Education have the least effect on treatment assignment.

Contrasting with the Random Forest and Gradient Boosting Machine models, the Uplift

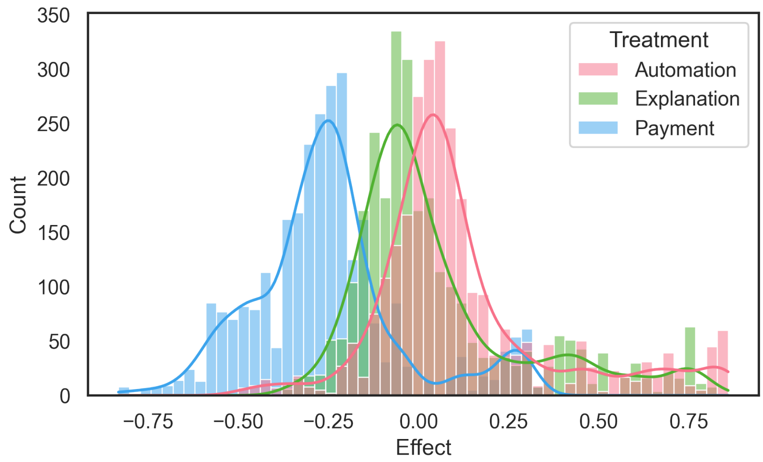


Figure 5.5: Distribution of Predicted Treatment Effects (Uplifts)

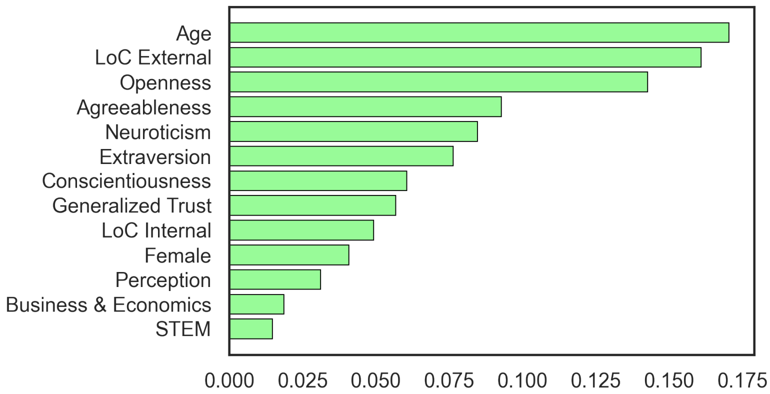


Figure 5.6: Feature Importances - Causal Model

model emphasizes the impact of these variables on the treatment effects rather than the outcome itself. While age and certain personality traits like Extraversion and Openness are influential across all models, the Uplift model uniquely demonstrates their role in optimizing treatments for delegation.

Evaluating causal inference models, like uplift random forests, is intricate due to counterfactual outcomes. We can only observe a given individual's delegation decision under one treatment. Unlike traditional classification, where predicted outcomes are compared to observed labels (as in table 5.4.5), uplift modeling predicts the difference between observed and unobserved counterfactual outcomes. This lack of observed outcomes for both scenarios for an individual restricts using standard classification metrics. Instead, metrics specific to uplift models, such as uplift curves, assess their performance. The uplift curve, similar in interpretation to the ROC curve, plots cumulative gain from targeting individuals by predicted uplift. Derived from it, the Area Under the Uplift Curve (AUUC) mirrors the AUC-ROC, gauging the model's ability to prioritize effective interventions. Figure 5.7 shows our model's Uplift Curve.

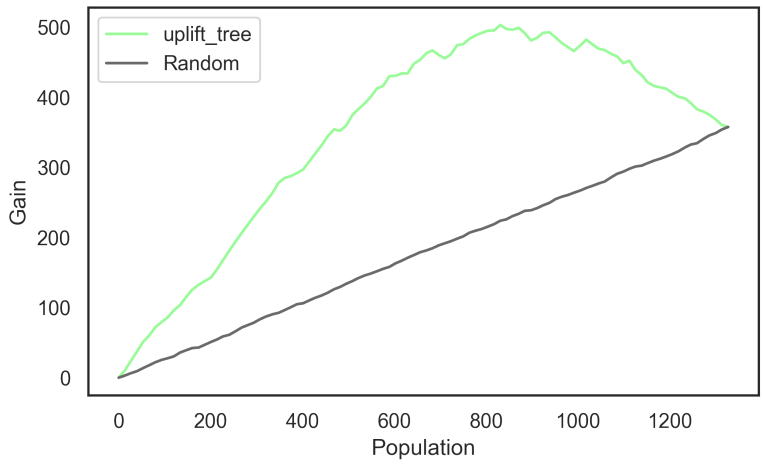


Figure 5.7: Uplift Curve

We have computed the AUUC using a synthetic control group consisting of individuals whose predicted optimal treatment matches the actual treatment they received or those in the actual control group, following the method in H. Chen, Harinen, Lee, Yung, and Zhao (2020). The uplift score for each individual in the synthetic control was computed, and individuals were ranked based on these scores. The AUUC was then calculated as the area under the curve plotting the cumulative proportion of actual outcomes against the proportion of the population targeted. The result is 0.977, which indicates relatively high performance in the prediction task and in explanation power.

Applying Uplift Random Forest to our study has offered valuable insights into the factors that influence the impact of treatments in delegation decisions. The model identified age, openness, and certain personality traits as significant determinants. It provided an additional perspective by focusing on the influence of these variables on treatment effectiveness rather than on the outcome itself.

5.4.7 How Subjects React to Non-Profitable Algorithmic Decisions

Numerous studies show that people initially trust algorithms, but trust may plummet after a mistake occurs (Glikson & Woolley, 2020). Dietvorst et al. (2015) found that people avoid algorithms or computerized decision-making systems even if they make fewer errors than humans due to high expectations for algorithms and attributing errors solely to the algorithm. Prahl and Van Swol (2017) showed that people are less likely to follow advice from a computer algorithm immediately after receiving incorrect advice. Complementarily, Chong, Zhang, Goucher-Lambert, Kotovsky, and Cagan (2022) reveals that poor algorithmic performance harms human confidence in the algorithm and self-confidence. Bogert et al. (2021) complements the idea of adverse reactions by outlining that bad decisions generated by algorithms are more severely punished than those of humans. To investigate this further, we analyzed participants' reactions after delegating a decision to the algorithm and receiving no payoff.

Delving into the impact of the algorithms' performance on the subjects, we calculated the frequency of participants changing their strategies from "delegate" to "not delegate" relative to the number of times the algorithm's decision resulted in a zero payoff, which does not necessarily mean a "wrong" choice but can also indicate a non-realized payoff from the "correct" choice. We extended this analysis to explore potential gender effects. Table 5.4.7 presents the absolute proportions of reaction results categorized by gender and treatment.

	Baseline	Explanation	Payment	Automation
General (aggregated)	0.30	0.25	0.35	0.09
Males	0.26	0.15	0.27	0.07
Females	0.34	0.31	0.40	0.10

Table 5.6: Relative frequencies of changing strategies (reaction) following algorithmic failures

On average, participants in the payment treatment group exhibited the highest reaction frequency (0.35), suggesting that individuals are more likely to change their decision when a financial incentive is involved. Conversely, the automation treatment group had the lowest frequency of reaction (0.09), indicating that participants are less likely to change their decision when the task is automated, possibly due to the complete handover process or also satisfaction with the algorithm performance, which was overall higher in the automation treatment (further details on the algorithm's performance are documented in chapter 5.4.8).

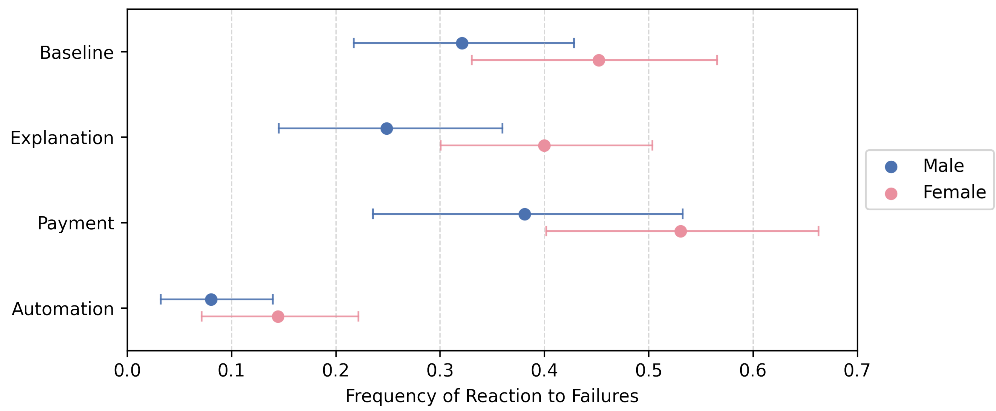


Figure 5.8: Frequencies of reaction to algorithmic failures by treatment and gender

Comparing reaction frequencies between males and females reveals that females have a higher reaction frequency across all treatments, suggesting they might be more sensitive to algorithm mistakes (figure 5.8). To further examine the gender gap in reaction, given that gender differences were not observed elsewhere in the experiment, we conducted statistical tests on both samples.

Similar to the statistical tests performed on the relative frequencies of delegation, we calculated the relative frequencies of reaction for each participant over 40 periods, treating each participant's decision path as an independent observation and separating the samples by gender. We then applied a Mann-Whitney U test (Mann & Whitney, 1947) to measure the difference between the two independent samples. The results show a value of 7751.51 and a p-value of 0.0028, outlining a statistically significant difference between the means of the frequency of strategy reactions for males and females. To deepen our understanding of participant reactions, we further analyzed whether contextual or personal factors influenced their behavior. Similar to the methodology used in the delegation behavior analysis (Chapter 5.4.3), we employed a logistic regression with standard errors clustered at the participant level. The results of this analysis are compiled in Table 5.7.

Variable	Coefficient	Standard Error	p-value
Constant	-2.188	0.868	$^{*}0.012$
Explanation	-0.088	0.181	0.626
Payment	-0.32	0.214	0.134
Automation	-1.042	0.24	$^{***} < 0.001$
Female	0.453	0.154	$^{**}0.003$
Age	0.005	0.016	0.775
STEM	-0.345	0.213	0.106
Business & Economics	0.021	0.182	0.908
Extraversion	0.031	0.053	0.567
Agreeableness	0.121	0.081	0.138
Conscientiousness	0.027	0.082	0.743
Neuroticism	-0.044	0.061	0.465
Openness	0.04	0.09	0.657
Internal LoC	-0.292	0.087	$^{***} < 0.001$
External LoC	-0.107	0.104	0.304
Generalized Trust	-0.034	0.07	0.63
Perception	-0.229	0.122	0.06

Note: Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 5.7: Logistic Regression Results - Reactions

The analysis indicates that task automation, gender, and internal locus of control are key factors in strategy changes following unprofitable algorithm decisions. Full task automation and a high internal locus of control reduce the likelihood of strategy shifts, suggesting trust in the process and personal control beliefs. Conversely, female participants are more prone to strategy changes, hinting at potential gender differences in reactions to algorithmic failures. Other factors, including algorithm explanation, payment requirement, and various personality traits, don't significantly influence strategy changes, suggesting their impact may be less direct.

5.4.8 Task Performance and Human-Algorithm Interaction

Finally, to evaluate the performance of Reinforcement Learning in the product selection task, we analyzed the mean probabilities of selecting each product quality level, grouping them based on their probabilities of receiving a payoff. The task was not straightforward due to the possibility of receiving a zero payoff even after identifying the best option, which could alter the weight of correct attractions. This ambiguity challenged human subjects and affected the algorithms' convergence capabilities. Figure 5.9 illustrates the development of choice probabilities for each product type.

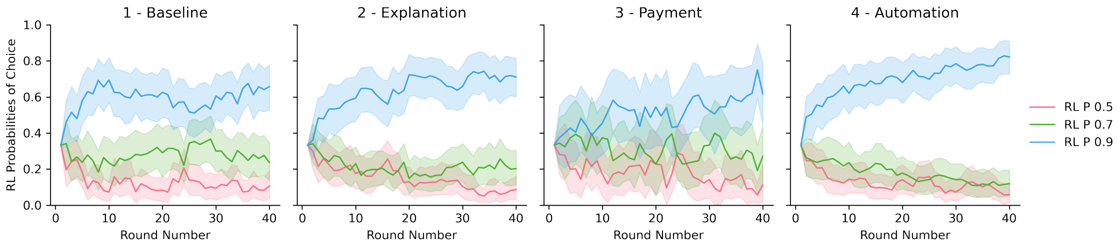


Figure 5.9: RL Choice Probabilities Over Time

In all instances, the algorithm could identify the highest quality product compared to the inferior alternatives. However, performance levels varied across treatments. We observed improved performance in generating optimal choice probabilities in treatments with higher delegation rates, such as explanation and automation, compared to the other groups, with the payment group being the most impacted. In treatments where participants exhibited higher "trust" in the algorithmic decision-making process, the performance in identifying the optimal product was better.

	Algorithm	Human
Baseline	0.592	0.511
Explanation	0.634	0.495
Payment	0.506	0.505
Automation	0.694	0.515

Table 5.8: Frequency "high" Product Selected

In a complementary analysis, table 5.4.8 compares performance between the algorithm and human subjects throughout the task. The values denote the success frequencies, normalized by the number of human or algorithm decisions. As expected, even with a non-trained algorithm that learned on the spot, the algorithm consistently outperformed the human subjects.

5.5 Conclusion and Discussion

This paper investigated the impact of framing conditions, explainability, willingness to pay, and complete task automation on delegation behavior in the context of algorithmic decision-making. Additionally, the study explored individual differences by examining the Big Five Personality Traits, Locus of Control, Generalized Trust, and other individual characteristics such as gender, age, and education.

We investigated the algorithm aversion phenomenon employing a multi-stage analysis covering hypotheses testing, regressions, machine learning, and causal inference models. Our findings revealed that context conditions significantly influenced participants' delegation choices. The study demonstrated that explaining the algorithm used in the product selection task improved user trust and increased the likelihood of delegation. In contrast, introducing a cost for delegation (Payment treatment) led to a decrease in delegation rates. Finally, the Automation treatment highlighted that participants were likelier to delegate decisions to the algorithm when the task was wholly automated.

In the machine learning application, we adopted a two-pronged approach to decipher the complex dynamics of decision delegation. We utilized traditional machine learning models — Random Forest and Gradient Boosting Machines — and an Uplift Random Forest model, providing complementary perspectives on the influences on delegation behavior. The Random Forest and Gradient Boosting Machine models offered insights into the direct impacts of individual and contextual variables on delegation decisions. Age, personality traits like Neuroticism and Extraversion, and factors like Payment and Perception consistently emerged as significant influences. These models underscored the intricate interplay of individual traits and contextual conditions, with no single factor dominating the decision to delegate. Complementing this, our Uplift Random Forest model provided direct heterogeneous treatment effects, which confirmed the impacts observed in the statistical analysis: the strong negative influence from the payment context and the moderately strong positive influence of automation. As for the impact of personal variables, this focused on their influence on quantifying the likelihood of delegation. Key variables such as Age, Openness, and certain personality traits significantly shaped the uplifts in treatment assignments. The model highlighted the importance of these factors in optimizing interventions to enhance delegation, adding a unique dimension to our understanding. The machine learning analysis revealed a nuanced understanding of how individual characteristics and contextual factors, alongside their interplay, shape decision delegation to algorithms. The machine learning models identified a set of influential factors with high predictive accuracy, while the Uplift model shed light on optimizing intervention impacts. This complexity and interconnectedness of personal and contextual factors were also reported by Snijders, Conijn, de Fouw, and van Berlo (2023). These insights provide valuable guidance for practitioners designing algorithmic decision systems, emphasizing the need for a personalized, context-sensitive approach.

In examining responses to algorithmic errors, we discovered pronounced reactions in scenarios involving payment treatments. Interestingly, these reactions were significantly more frequent among females, indicating the presence of gender effects. Confirmatory statistical analyses reinforced these observations, revealing that factors such as Automation, Payment, and Internal Locus of Control significantly influenced participants' responses to algorithmic mistakes. These findings highlight the influence of both gender and specific situational contexts and confirm previous experiments in the literature pointing to algorithmic failures as a driver of aversion.

Employing a non-biased algorithm and allowing it to learn exclusively from the interaction with participants allowed us to observe how the algorithm's learning process evolved alongside the participants' decision-making behavior. In particular, treatments with lower delegation rates negatively affected the algorithm's performance, generating sub-optimal choice probabilities.

The implications of these findings are manifold. By better understanding the factors influencing delegation behavior in algorithmic decision-making, we can develop more user-friendly systems that facilitate trust and encourage appropriate delegation. These insights can contribute to designing decision support tools tailored to individual preferences and optimize human-algorithm collaboration.

This study has several limitations, including the simplicity of the experimental design, which may not fully capture the complexity of real-world decision-making scenarios, and a potentially non-representative sample. The interconnectedness and multicollinearity of personal traits also present challenges in isolating and interpreting their individual effects on delegation behavior. Further research could employ more realistic product designs and decision-making tasks and investigate the effects of combined treatment conditions, e.g., payment and automation, payment and explanation, among other things, to understand better the interplay between various contextual factors and their impact on delegation behavior in algorithmic decision-making. Moreover, future studies could also consider evaluating purely economic behaviors and attitudes, such as risk, loss, and ambiguity aversion.

In conclusion, this paper contributes to the growing literature on algorithm aversion and delegation behavior. It highlights the importance of framing conditions, explainability, individual differences, and the complex interaction between variables in shaping user preferences and trust in algorithmic decision-making systems. Future research could delve deeper into the interaction between these factors and explore the impact of different explanation styles, varying costs for delegation, and other contextual factors on delegation behavior. By understanding the nuances of human-algorithm collaboration, we can develop systems that enhance decision-making and contribute to more efficient and effective outcomes in various domains.

5.6 Appendix

5.6.1 Methodological Formalizations

This chapter provides an overview of the machine learning methods used in the project. The following subchapters account for the Random Forest, Gradient Boosting, and Uplift Random Forest methods, providing generalizations of the algorithms' implementations.

Reinforcement Learning Implementation and Tuning

The underlying problem introduces three options or products, expressed as Q_i, each associated with distinct probabilities of receiving a payoff that can be selected at each period, t. Each product Q_i is associated with an attraction value $A_{Q_i}(t)$, representing the decision weight attached to product Q_i at period t. Following the theoretical frameworks in

C. Camerer and Hua Ho (1999); Erev and Roth (1998), the attraction values are updated based on the payoffs received by selecting product Q_i using the following update rule:

$$A_{Q_i}(t) = \phi A_{Q_i}(t-1) + I(Q(t) = Q_i)\pi_{Q_i}(t) \tag{5.1}$$

This model features the indicator function, which means that a player's attraction to a strategy can only increase if they choose it. The attraction increases by the amount of payoff received from it. In the update rule, the indicator functions $I(Q(t) = Q_i)$ equals 1 if a participant chooses product Q_i at period t and 0 otherwise, while $\pi_{Q_i}(t)$ represents the payoff received when choosing product Q_i at period t. The recency parameter ϕ indicates how quickly past payoffs are forgotten, which acts as a form of learning rate. Attractions from the previous period determine choice probabilities in any period. A logistic transformation over the attraction values calculates the probabilities:

$$P_{Q_i}(t+1) = \frac{e^{\lambda A_{Q_i}(t)}}{\sum_{k=1}^{m} e^{\lambda A_{Q_k}(t)}} \tag{5.2}$$

In this equation, $P_{Q_i}(t+1)$ represents the probability of selecting product Q_i at time $t+1$, $A_{Q_i}(t)$ denotes the attraction of product Q_i at time t, and m indicates the number of available product options. The second parameter, λ, reflects the sensitivity of choice probabilities to differences in attractions. The two necessary parameters were tuned using observed data from 1000 simulations, testing for the ranges $0-1$ for ϕ and $0-10$ for λ. The tuning resulted in $\phi = 0.47$ and $\lambda = 4.5$, associated with higher payoffs. The experiment parameters were set to these values statically.

Random Forest

The Random Forest algorithm concept builds a large collection of de-correlated decision trees and then aggregates them through a majority voting system for classification problems. Hastie et al. (2009) generalized the algorithm as follows:

More details on the Random Forest algorithm can be found in Breiman (2001).

Gradient Boosting Machines

Gradient Boosting Machines (GBM) is a machine learning method that builds a sequence of decision trees, each correcting its predecessor's mistakes, to create a final, robust predictive model (J. H. Friedman, 2001). Hastie et al. (2009) also provides a generalization of this model, with the stepwise algorithm defined as:

Lines 2-6 are repeated K times at each iteration m, once for each class. For a more detailed description of the Gradient Boosting Machines and their derivations, see the comprehensive overview in Hastie et al. (2009).

Uplift Modelling

The underlying method is the same as that of the Random Forest. However, For the uplift random forest classifier, the uplift tree consists of a combination of methods based on uplift

Algorithm 3 Random Forest Algorithm

Require: B trees to be grown, N size of bootstrap sample, M total variables, m selected variables, $n_{\min}$ minimum node size

Ensure: Output the ensemble of trees $\{T_b\}_1^B$

```
for b = 1 to B do
    Draw a bootstrap sample of size N from the training data
    Grow a decision tree T_b on this data by:
    while each terminal node of the tree until the minimum node size n_min is reached do
        Select m variables at random from all M variables
        Pick the best variable/split-point among the m
        Split the node into two daughter nodes
    end while
end for
To make a prediction for a new point x, let Ĉ_b(x) be the class prediction of the bth random forest tree
The random forest chooses Ĉ_rf(x) = majority vote{Ĉ_b(x)}_1^B
```

Algorithm 4 Gradient Boosting Machines Algorithm (Generalized)

Require: M iterations, n number of observations, L loss function, y_i observed response, $F(x_i)$ predicted response, $h_m(x)$ base learner at iteration m

Ensure: Output $F_M(x)$ as the final model

Initialize the model with a constant value:

$$F_0(x) = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n} L(y_i, \gamma)$$

for $m = 1$ to M **do**

Compute pseudo-residuals:

$$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)}, \quad \text{for}\, i = 1, \ldots, N.$$

Fit a base learner $h_m(x)$ to pseudo-residuals, i.e., train it using the training set $\{(x_i, r_{im})\}_{i=1}^n$

Compute multiplier:

$$\gamma_j m = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i))$$

Update the model:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$$

end for

modeling, with the tree split criterion based on differences in the uplift. In the standard notation (Rubin, 1974), we consider $Y_i(1)$ an individual's i being treated and $Y_i(0)$ for being in the control group. In this case, the causal effect τ_i is given by $\tau_i = Y_i(1) - Y_i(0)$. Having $W_i \in 0, 1$ as a binary variable indicating if person i is in the active treatment group, and 0 otherwise (control group), the observed outcome is $Y_i^{obs} = W_i Y_i(1) + (1 - W_i) Y_i(0)$.

Based on Gutierrez and Gérardy (2017), considering a balanced, randomized experiment, the average treatment effects (uplifts) are estimated as:

$$\hat{\tau} = \underbrace{\frac{\sum_i Y_i^{obs} W_i}{\sum_i W_i}}_{p} - \underbrace{\frac{\sum_i Y_i^{obs}(1 - W_i)}{\sum_i (1 - W_i)}}_{q}, \tag{5.3}$$

Which represents the difference in the sample average outcome between the treated and untreated observations. For the splitting criterion, the gain difference after splitting is defined as:

$$D_{gain} = D_{after_{split}}(P^T, P^C) - D_{before_{split}}(P^T, P^C) \tag{5.4}$$

Where D is the difference and P^T and P^C is the probability distribution of the outcome variable in the treatment and control groups (Rzepakowski & Jaroszewicz, 2012). The uplift trees were split using the Chi function, rooted in a statistical test that determines significant associations between two categorical variables. Within uplift modeling, this function aids in prioritizing splits that highlight a significant relationship between the treatment and the outcome. The divergence in this method is represented by X^2:

$$X^2(P : Q) = \sum_{k=left,right} \frac{(p_k - q_l)^2}{q_k} \tag{5.5}$$

where p indicates the sample mean in the treatment group, q is the sample mean in the control group, and k denotes the leaf in which p and q are calculated.

5.6.2 Additional Data and Analyses

This chapter presents additional data analysis elements not included in the main manuscript.

Correlations

Delegation behavior exhibits weak positive correlations with STEM degrees, extraversion, agreeableness, conscientiousness, internal locus of control, and external locus of control. Conversely, it has weak negative correlations with gender (female), business and economics degrees, and neuroticism. Age and openness display almost no correlation with delegation behavior (figure 5.10).

Table 5.6.2 displays the results of point-biserial correlation coefficients between the personality traits and delegation behavior (binary).

Regressions

This regression model includes interaction terms to account for the correlation between independent variables (table 5.6.2), providing a more nuanced analysis of the relationships between variables and delegation behavior. In this model, the main effects of some variables

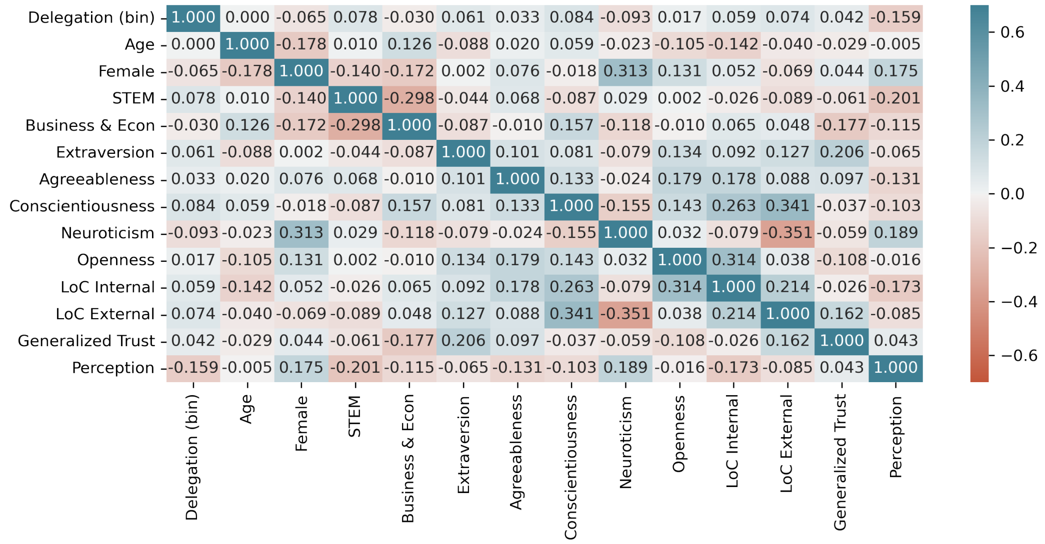


Figure 5.10: Spearman correlation coefficients

change, and the added interaction terms help us better understand how the relationships between variables affect the outcome.

The internal locus of control variable becomes significant ($p = 0.041$) in the model with interaction terms, while it was not significant in the model without interactions. This change suggests that the relationship between internal locus of control and delegation behavior might be more complex than initially estimated by the first model. Including interaction terms allow us to capture the combined effects of internal locus of control with other variables, such as openness, which might help explain this shift in statistical significance.

The interaction between female gender and neuroticism is significant at the 10% level ($p = 0.084$). For instance, women generally report higher neuroticism scores than men (Costa Jr, Terracciano, & McCrae, 2001; Schmitt, Realo, Voracek, & Allik, 2008; Weisberg, DeYoung, & Hirsh, 2011), which is also true for our sample. Given that women generally report higher neuroticism scores than men, this term indicates that the relationship between neuroticism and delegation behavior differs for males and females. Specifically, the effect of neuroticism on delegation behavior may be more substantial for one gender than the other. As a result, the positive coefficient for the female gender in the second model suggests that the likelihood of delegation among females might depend more on their neuroticism level than males.

Another noteworthy interaction term is the one between internal locus of control and openness, which is significant at the 10% level ($p = 0.080$). This interaction suggests that the effect of internal locus of control on delegation behavior is more pronounced for individuals with specific levels of openness. For example, participants with a high internal locus of control and high openness might be more likely to delegate tasks than those with a high internal locus of control and low openness. This finding further emphasizes

Variable	Correlation Coefficient	p-value
Age	-0.019	$^{*}0.026$
Female	-0.065	$^{***} < 0.001$
STEM	0.078	$^{***} < 0.001$
Business & Economics	-0.03	$^{***} < 0.001$
Extraversion	0.06	$^{***} < 0.001$
Agreeableness	0.039	$^{***} < 0.001$
Conscientiousness	0.089	$^{***} < 0.001$
Neuroticism	-0.087	$^{***} < 0.001$
Openness	0.024	$^{**}0.005$
Internal LoC	0.06	$^{***} < 0.001$
External LoC	0.068	$^{***} < 0.001$
Generalized Trust	0.044	$^{***} < 0.001$
Perception	-0.147	$^{***} < 0.001$

Note: Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 5.9: Point-biserial correlation coefficients to binary action of delegation

Variable	Coefficient	Standard Error	p-value
Constant	5.195	4.046	0.199
Explanation	0.195	0.21	0.354
Payment	-1.053	0.24	$^{***} < 0.001$
Automation	0.453	0.235	0.054
Female	0.821	0.567	0.147
Age	-0.012	0.018	0.501
STEM	0.269	0.232	0.246
Business & Economics	-0.193	0.201	0.337
Extraversion	0.018	0.059	0.755
Agreeableness	0.044	0.073	0.552
Conscientiousness	-0.411	0.597	0.491
Neuroticism	0.111	0.361	0.759
Openness	-0.71	0.416	0.088
Internal LoC	-1.266	0.618	$^{*}0.041$
External LoC	0.251	0.604	0.678
Generalized Trust	0.07	0.065	0.284
Perception	-0.361	0.14	$^{*}0.01$
Female x Neuroticism	-0.234	0.136	0.084
Internal Loc x Conscientiousness	0.136	0.089	0.129
External Loc x Conscientiousness	-0.043	0.094	0.65
External Loc x Neuroticism	-0.011	0.079	0.894
Internal Loc x Openness	0.139	0.079	0.08

Note: Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 5.10: Logistic Regression results - delegation, with interaction Terms

the importance of considering the interaction effects when examining the relationships between variables and delegation behavior.

We also have fit quantile regression models (Koenker & Bassett Jr, 1978) using the cumulative frequency of delegation for each participant across all periods, removing the time dimension. We employed this method due to the varying relationships between the variables across different parts of the outcome distribution and the lack of normality. The results are summarized in table 5.6.2.

Variable	Coefficient	Standard Error	p-value
Intercept	0.386	0.28	0.169
Explanation	0.088	0.059	0.138
Payment	-0.367	0.06	$^{***} < 0.001$
Automation	0.192	0.06	$^{**}0.002$
Female	-0.027	0.046	0.558
Age	-0.003	0.005	0.526
STEM	0.078	0.057	0.173
Business & Economics	-0.051	0.055	0.348
Extraversion	0.023	0.017	0.179
Agreeableness	0.011	0.021	0.598
Conscientiousness	0.025	0.024	0.285
Neuroticism	0.002	0.018	0.927
Openness	0.005	0.023	0.841
Internal LoC	-0.02	0.027	0.467
External LoC	0.058	0.028	$^{*}0.04$
Generalized Trust	0.013	0.019	0.49
Perception	-0.143	0.034	$^{***} < 0.001$

Note: Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 5.11: Quantile Regression results - cumulative delegation frequencies

This model explains approximately 18.95% of the sample variance. Similarly to the logistic regression results, these findings show that the condition involving payment significantly reduces the frequency of delegation ($p < 0.001$), while full automation significantly increases it ($p = 0.002$). Among personal characteristics, only External Locus of Control significantly contributes to delegation, indicating that participants who believe outcomes are beyond their control are more likely to delegate decisions ($p = 0.04$). Moreover, a negative perception of algorithms significantly corresponds to a less frequent delegation of decisions ($p < 0.001$). Other actors such as explanation condition, demographics, Big Five personality traits, Internal Locus of Control, and Trust do not significantly affect the delegation frequency. We have also controlled for correlated variables in this model by adding interaction terms; the results are summarized in table 5.6.2.

Upon adding interaction terms, the pseudo-R-squared value rose to 21.01%, showing a marginally improved model fit. Payment ($p < 0.001$) and automation ($p = 0.01$) still significantly influence delegation. Notably, individuals with a STEM background ($p = 0.017$) show a significant positive association with delegation. Openness to experience negatively correlates with delegation ($p = 0.034$). A significant interaction emerges between internal locus of control and Openness ($p = 0.04$): those high in internal locus of control and openness tend to delegate more. A negative view of algorithms remains a strong deterrent

Variable	Coefficient	Standard Error	p-value
Intercept	1.206	1.043	0.248
Explanation	0.039	0.057	0.497
Payment	-0.397	0.058	$^{***} < 0.001$
Automation	0.151	0.058	*0.01
Female	0.105	0.139	0.45
Age	-0.007	0.005	0.14
STEM	0.131	0.055	*0.017
Business & Economics	-0.034	0.052	0.512
Extraversion	0.011	0.017	0.509
Agreeableness	0.015	0.02	0.451
Conscientiousness	0.058	0.166	0.727
Neuroticism	0.02	0.086	0.814
Openness	-0.22	0.103	*0.034
Internal LoC	-0.258	0.144	0.074
External LoC	0.18	0.157	0.253
Generalized Trust	0.008	0.018	0.672
Perception	-0.102	0.033	**0.002
Female x Neuroticism	-0.032	0.033	0.331
Internal Loc x Conscientiousness	0.016	0.026	0.548
External Loc x Conscientiousness	-0.025	0.025	0.307
External Loc x Neuroticism	-0.003	0.018	0.879
Internal Loc x Openness	0.041	0.02	*0.04

Note: Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 5.12: Quantile Regression Results - cumulative delegation frequencies, with interaction terms

to delegation ($p = 0.002$).

5.6.3 Technical Remarks

The documented experiment was executed online, programmed with the oTree open-source platform (D. L. Chen et al., 2016). The data work was performed using Python language. The statistical tests were done using statsmodels (Seabold & Perktold, 2010). The machine learning models were deployed, tuned, and cross-validated using Scikit-Learn (Pedregosa et al., 2011). Both models were tuned using a grid search algorithm with the target to maximize the AUC-ROC. It is important to outline that this is a computationally expensive procedure. The parameter set for the Random Forest model is in table 5.13.

Similarly, the grid search-generated parameters for the GBM model are described in table 5.14

The cross-validation technique used in both models was the GroupKFold algorithm, which aggregated samples for the same participant. This procedure was performed in both the parameter search and model training steps, using five validation folds.

The uplift random forest classifier was implemented using the causalml library (H. Chen et al., 2020). Since this method, in conjunction with the group cross-validation using syn-

Parameter	Value	Definition
bootstrap	True	Determines whether or not to use bootstrap samples when building trees
class_weight	balanced_subsample	Adjusts the weights of the classes. balanced_subsample means it computes weights based on the bootstrap sample for every tree
criterion	entropy	Defines the function to measure the quality of a split. entropy is for information gain
max_depth	15	Specifies the maximum depth of the tree
max_features	auto	The number of features to consider when looking for the best split. auto means the square root of the total number of features
min_samples_leaf	1	The minimum number of samples required to be at a leaf node
min_samples_split	min_samples_split	The minimum number of samples required to split an internal node
n_estimators	100	The number of trees in the forest

Table 5.13: Random Forest Classifier parameters

Parameter	Value	Definition
learning_rate	0.05	Determines the impact of each tree on the final outcome
max_depth	10	Specifies the maximum depth of the tree
max_features	sqrt	The number of features to consider when looking for the best split. sqrt means the square root of the total number of features
min_samples_leaf	1	The minimum number of samples required to be at a leaf node
min_samples_split	15	The minimum number of samples required to split an internal node
n_estimators	100	The number of boosting stages to perform. Each stage adds a new tree into the ensemble
subsample	0.7	The fraction of samples to be used for fitting the individual base learners

Table 5.14: Gradient Boosting Machine Classifier parameters

thetic control groups, was performance costly, we implemented a less-exhaustive approach for the parameter-fitting method, using the Optuna library (Akiba et al., 2019). It employs efficient search algorithms, such as Tree-structured Parzen Estimator (TPE). We ran an optimization study for 150 trials and selected the parameter set that yielded satisfactory AUUC scores. One important remark here is that calculating the AUUC in this way might produce abnormally high results due to the stochastics in place, so practitioners might have to supervise the optimization process. Table 5.15 describes the parameter values.

Parameter	Value	Definition
n_estimators	850	The number of trees in the forest
max_depth	8	The maximum depth of each decision tree
max_features	9	The number of features to consider when looking for the best split
min_samples_leaf	45	The minimum number of samples required to be at a leaf node
min_samples_treatment	15	The minimum number of samples in a leaf node that come from the treatment group
n_reg	14	The regularization parameter used in the causal tree procedure
evaluationFunction	Chi	The evaluation function used to evaluate splits

Table 5.15: Uplift Random Forest Classifier parameters

5.6.4 Experiment Design Screens

In this appendix session, we added the most important screens for the experiment. Figure 5.13 contains the main task screens for each treatment. Figure 5.14 shows the attention questions.

5.11 Displays the information flowchart adopted in the experiment for the different treatments, where the automation treatment has a dedicated loop

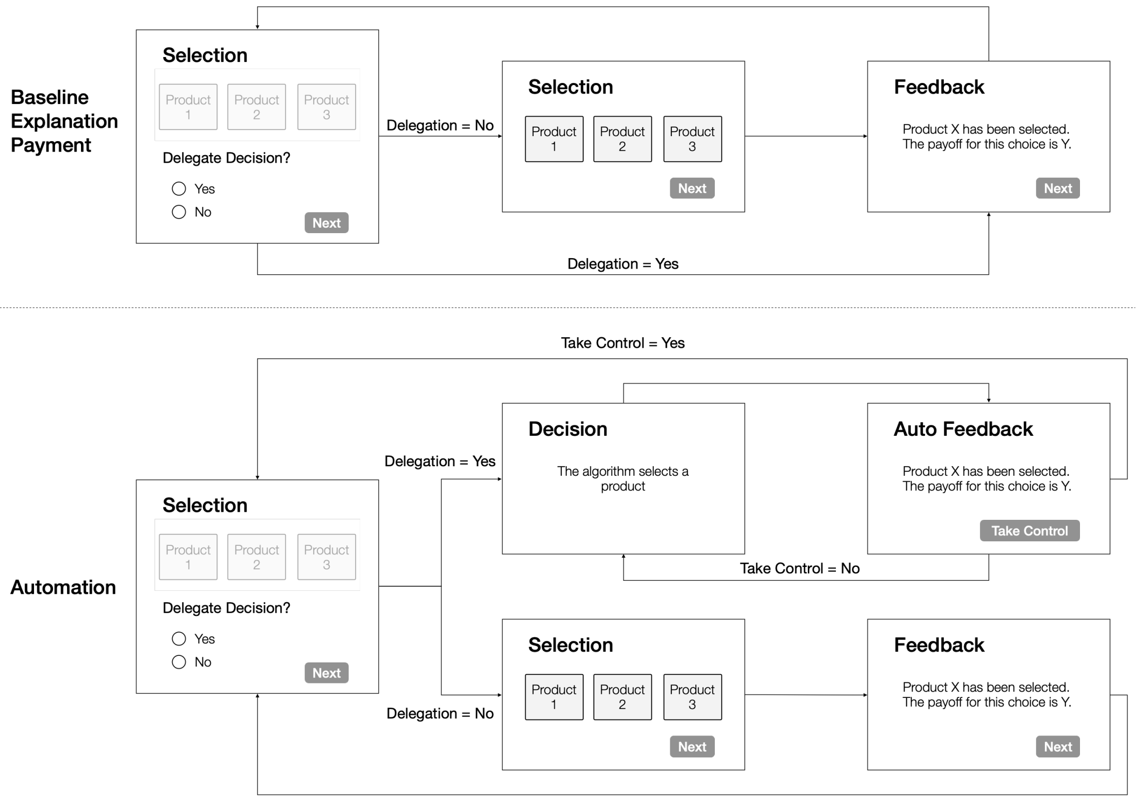


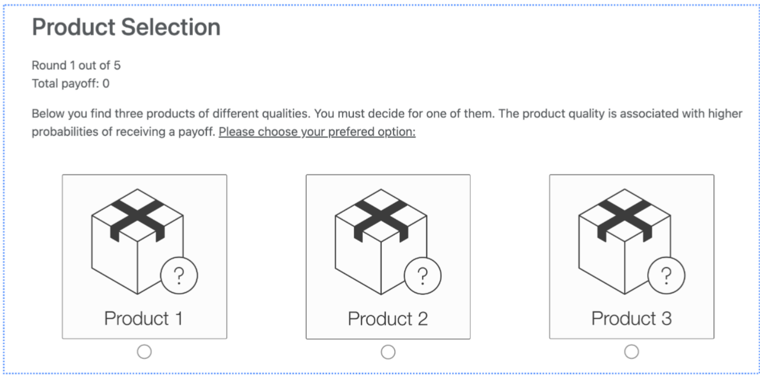
Figure 5.11: Experiment Flowchart

Instructions

In this experiment, you will be presented with three product choices, which will be named generically.

Product Selection

Each of the products has a **quality level**, which will be randomly allocated between the three alternatives. During the experiment, you will have to select one of the products, and you will receive a payoff according to your selection, as in the example below:



The product quality levels introduce different probabilities of acquiring points as a payoff, defined as:

- **High quality** – 0.9 probability of receiving 1 point
- **Medium quality** – 0.7 probability of receiving 1 point
- **Low quality –** 0.5 probability of receiving 1 point

You will have the opportunity to perform repeated product selections for 40 rounds. The product qualities will remain unchanged across the rounds after initial random allocation. For example, if your quality allocation was Product 1: High, Product 2: Low, and Product 3: Medium; these values will remain throughout the experiment.

If you learn which product is associated with which quality level, you will have higher chances of better payoffs.

Decision Delegation

In each round, you will have the opportunity to delegate your decision to an algorithm, which will select the product for you. If you choose to delegate, you get feedback on the selection of the algorithm and move to the next round. If you do not delegate, you can make the product selection yourself.

Questionnaire

When the experiment is finished, you will be redirected to a questionnaire with simple multiple choice questions. After that, you will the conclusion page, where your results will be shown, as well as the final results and your earnings.

To start with the experiment, please click "next".

Figure 5.12: General Instructions Screen

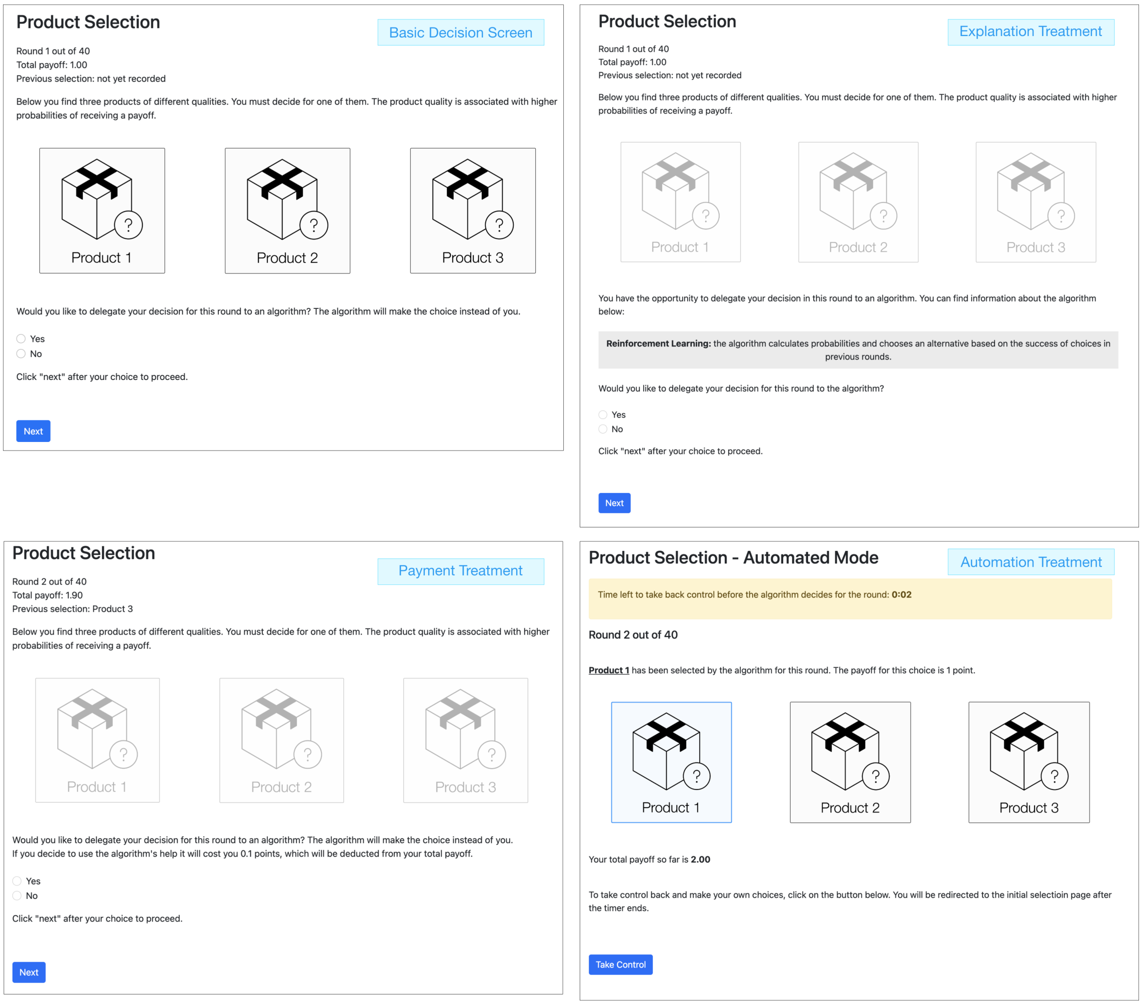


Figure 5.13: Main Task Experiment Screens

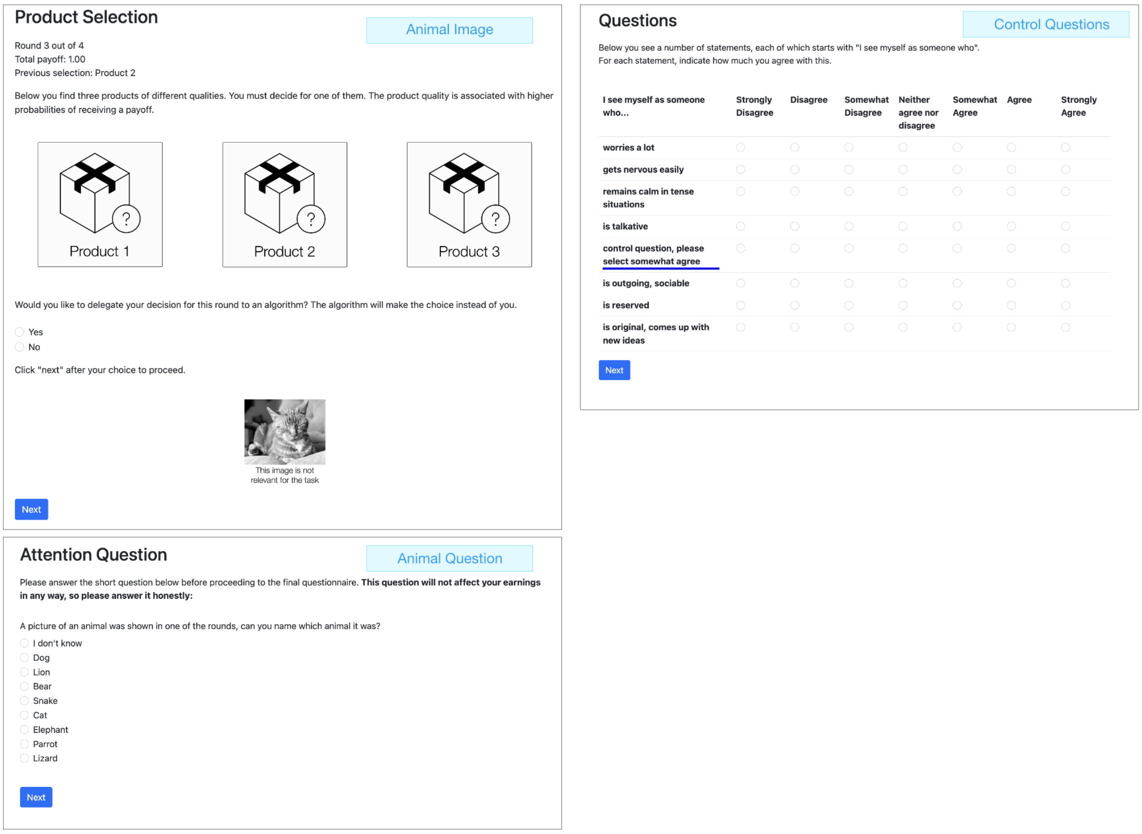


Figure 5.14: Attention Measures

6. Motivated Sampling of Information: Analysis With Experimental Data and Agent-Based Modeling in a Bayesian Framework

Authors

Vinícius Ferraz & Leon Houf

Abstract

This paper investigates information sampling where objective and subjective criteria are coupled. This experimental framework creates a situation that gives room for motivated reasoning, which we identify as motivated sampling. We present participants with a binary sampling and decision task. Participants sample information from two "computers," which generate numbers from distinct distributions, and participants have to identify the "high distribution" computer. In this task, we vary externalities on the participants' decision to induce subjective preferences. Furthermore, we vary the type of feedback participants receive. Following this methodological framework, we compare participants with simulated agents using Reinforcement Learning variations. We find motivated reasoning in several instances. First, we show that female subjects sample significantly more than male subjects when faced with a negative externality or Bayesian posterior feedback. Moreover, we show a strong, intensive margin of motivated sampling. Here, subjects sample additionally from the option with positive externality if they deem it correct, which shows an added liking to sample from it. These findings provide an understanding of motivated sampling and a specific application of motivated reasoning, emphasizing the importance of subjective preferences, feedback, and gender differences in all situations where information sampling is necessary for decision-making.

Keywords

Motivated Reasoning, Information Sampling, Bayesian Learning, Decision Theory

6.1 Introduction

Decision-making is an integral part of human life. Individuals are frequently faced with the task of selecting from multiple options, whether choosing a restaurant, booking travel tickets, or picking a university. In these situations, information sampling is crucial to facilitate well-informed choices. Understanding the dynamics of information sampling can lead to better decision-making, especially in complex scenarios. Previous research has provided valuable insights into information sampling behavior and its cognitive and computational costs for human subjects (Kool & Botvinick, 2018; Petitet, Attaallah, Manohar, & Husain, 2021), individual factors in sampling and information-seeking (Gottlieb & Oudeyer, 2018; Kelly, Sharot, et al., 2021), decision-making perspectives (Leung, 2020) and how rewards influence sampling through a Pavlovian-approach (Hunt, Rutledge, Malalasekera, Kennerley, & Dolan, 2016).

However, there remain open questions about coupling subjective and objective criteria in decision-making. Specifically, how do individuals approach decisions when they aim for the objectively best option but already have a pre-existing preference? This phenomenon, known as *motivated reasoning*, is characterized by individuals processing information in a way that aligns with their pre-existing beliefs or desires (Bénabou & Tirole, 2016; Eil & Rao, 2011; Hagenbach & Koessler, 2022). This paper introduces the concept of *motivated sampling*. This phenomenon represents the overlap between information sampling and the well-established idea of motivated reasoning.

We present subjects with a binary decision task to disentangle the effects of objective and subjective criteria in information sampling. In this task, subjects must sample information to determine the objectively correct option to receive a payoff. We then asymmetrically add negative and positive externalities to the options. A positive externality is an additional reward for an organization the subject liked, while a negative externality is a reward for an organization the subject explicitly disliked. Through this, we induce subjective preferences into the sampling and decision situation. Using a between-subjects design, we can measure how these subjective criteria affect sampling behavior. In the simulations, we employ Reinforcement Learning Models for Optimal Sampling. The simulations give us a sampling behavior benchmark that helps us better understand human sampling.

Our central research question is: "How do subjective preferences on externalities influence motivated sampling?' Additionally, we analyze the accuracy of posterior beliefs with different feedback forms and the time participants actively engage in the task.

Our findings show that women sample significantly more information than men when a negative externality is at play. Subjects sample much more for the option with a positive externality when they deem this option correct than when incorrect. This behavior, termed motivated sampling, indicates a 'liking' to sample from options that meet both objective and subjective criteria. In contrast, we do not see such a behavior when a negative externality is involved. For both types of externality, male participants show a more substantial bias for the "nicer" option than female participants.

We offer a novel perspective on information sampling strategies by disentangling the effect of objective and subjective criteria. Specifically, we uncover the mechanisms of motivated sampling. Hereby, we add a specific application to the more broadly defined theme of motivated reasoning, also serving as a fundamental, underlying mechanism of confirmation bias.

The remainder of this paper is structured as follows: We first outline the method and experimental design. Then, we present the empirical results of our human subject experiment. Section four concludes.

6.2 Design and Methods

This section provides an overview of the experimental design used in our online study. We first characterize the participants' demographics. Then, we describe the main experimental task and provide an overview of the overall procedure. Next, we outline the treatment dimensions of our 3x3 between-subject design, allowing us to investigate the interplay between 3 forms of externalities and three forms of feedback. Lastly, we describe the simulation model dynamics.

6.2.1 Demographics

A total of 457 students were recruited from the experimental economics labs at Heidelberg University and Rhine-Waal University of Applied Sciences, both in Germany, with a 37% representation from Heidelberg and 63% from Rhine-Waal. The participants were split almost equally between male (48%) and female (51%), with 1% choosing not to disclose or identify as non-binary. The average age of the participants was 24.2 years old (standard deviation 4.6 years), and they represented a diverse mix of nationalities, with the largest group being German (48%) followed by Indian (10%). The remaining participants came from a variety of international backgrounds. The experiment was online, and participants could take part any time of the day or week and take breaks of any length. The median participant had the experiment open in the browser for 30.5 minutes. The average earning was 4.91€.

6.2.2 Description of Main Task

During the task, participants are presented with two "computers." Each computer generates numbers based on specific distributions, as depicted in figure 6.1. One of the computers produces higher numbers on average because the computer uses a "high distribution" of numbers, whereas the other computer uses a "low distribution." Both distributions produce numbers from 1 to 8, as used in Goette, Han, and Leung (2020). The "high computer" produces numbers using the distribution shown on the left side of figure 6.1, and the "low computer" produces numbers using the distribution on the right.

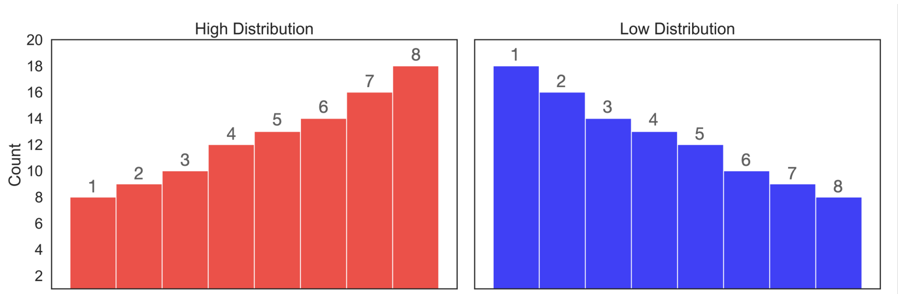


Figure 6.1: High and Low Distributions

In every round, one computer uses the high distribution of numbers, while the other uses the low distribution. The computer that uses the high distribution is determined randomly in each round with a 50/50 chance. The participants' goal is to identify which computer uses the high distribution. Each correct identification is rewarded with a point. Participants can sample as many numbers as they want by clicking on one of the computers, but they are restricted to only sample one new number every two seconds. This constraint is explained as a need for the computers to reload to produce the following number. We use this two-second restriction to prevent rapid ”over-clicking” by participants because we want to create a situation where every new information can be taken into account subsequently.

6.2.3 General Procedure of Experiment

Now, we describe the overall procedure of the experiment surrounding the main task. Before the experiment starts, we gather demographic information from participants, which includes two questions about which organizations participants would be *most* and *least* likely to contribute money to. Then, the experiment starts with a practice round and 20 payment-relevant rounds. Each round consists of three parts: Part A assesses the prior belief about which computer uses the high distribution, shown in figure 6.8 and 6.9. Part B is the main task explained above. Part C assesses the posterior belief after the participants chose their option, shown in figure 6.10. At the end of the 20 rounds, we randomly select three rounds as payoff relevant. For each point a participant scored in those rounds, they receive 1.5€, in addition to a 1.5€ show-up fee.

We have nine treatments that differ in whether an externality is added to the main task and whether feedback is provided after part C at the end of a round. Subjects are randomly assigned to their treatment group. The treatments are created as a 3×3 design along the dimensions of externality and feedback. And as pre-registered[1], we have 50-52 participants in every treatment cell as shown in table 6.1, allowing us to pool the cells across rows or columns when we compare the respective treatment dimension for feedback and externality.

		Feedback		
	Condition	Outcome	Bayes	No
Externality	Negative	51	50	52
	No	52	51	50
	Positive	50	51	50

Table 6.1: Subjects per Treatment

6.2.4 Externality Treatments

In the externality dimension, we distinguish between no externality, positive externality, and negative externality. With no externality, the main task is as described above. In the externality treatments, we attach an externality randomly to one of the two computers in each round.

In the positive externality treatments, we use the organization the participant chose in the demographics section as the organization they are *most* likely to give money to. This

[1] AsPredicted #104844

organization is attached to one of the computers. The organization receives 1 point in the round if the participant chooses the respective computer, and it is *correct* (shown in appendix figure 6.12). This feature introduces a subjective element to the decision task. While the objective criterion is still to select the correct computer, since only then, the participant *and* the preferred organization, receive a point, participants might have a subjective preference to experience the option with externality to be correct.

In the negative externality treatments, we use the organization the participant chose as *least* likely to give money to. This organization is attached to one of the computers, and the organization receives 1 point in the round if the participant chooses the respective computer, but that was the *wrong* decision (shown in appendix figure 6.13). Here, the objective incentive is still to select the correct computer, since then the participants receive a point, and the antagonizing organization receives nothing. The subjective element for the participants here is an increased subjective incentive not to be wrong when selecting the option with externality, as the antagonizing organization can only receive a point through a subject's mistake.

6.2.5 Feedback Treatments

In the feedback treatment dimension, we distinguish between no feedback, outcome feedback, and Bayes feedback. With no feedback, the procedure is as described above, and participants move to the next round without receiving any feedback.

In the outcome feedback treatments, participants learn whether their computer choice was correct or incorrect at the end of each round.

In the Bayes feedback treatments, at the end of each round, after stating their posterior belief in part C, participants receive a reminder of the posterior belief they just stated and are informed about the rational Bayesian posterior. The feedback is shown in appendix figure 6.15, and the calculation of the Bayesian posterior is outlined in appendix 6.5. The overall design is summarized in figure 6.2. Samples of the experiment screens for all tasks and treatment variants are documented in appendix 6.5.2.

6.2.6 Simulation Model

To model the motivated sampling problem in a computational framework, we wanted to understand how the sampling behavior unfolds by following robust decision rules. Especially when comparing both sampling and reward performances between human and artificial agents. Therefore, the introduced computational environment mimics the task of sampling for information and selecting a computer in each round while accounting for constraints imposed by sampling and externalities.

In the original problem's incentive schemes, sampling for numbers imposes a cost on the subjects translated to time since the sampling action button is disabled for 2 seconds after each click. The externalities, which could be either positive or negative, added an additional layer of incentives to the decision related to giving points to preferred (considered a gain) or disliked (considered a loss) organizations. Our simulation model accounts for

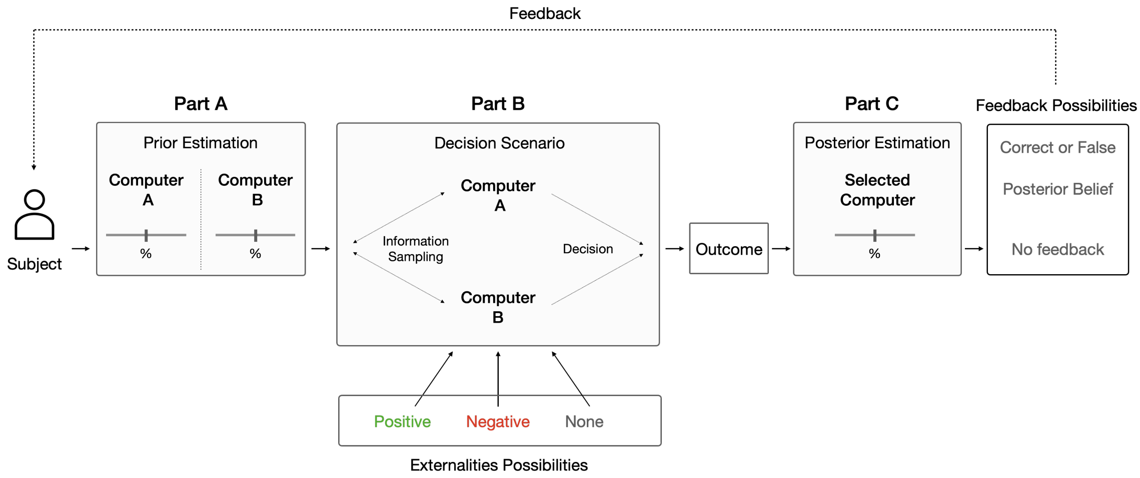


Figure 6.2: Experimental Design Framework

both constraints by deducting points from the reward in case of sampling actions or the realization of negative externalities and adding points in case of positive externalities. To model these conditions, we introduce this simple reward function: $R = P + E - C \times T$, where P is the payoff for the round, equal to the base payoff if the selected computer is correct or zero otherwise. E is the externality value added or subtracted from the reward depending on the externality condition. C is the cost per sample. T is the total number of samples. The externality component E of the reward follows the conditions: when the positive externality is realized, the participant gets additional points, and for the negative, the participant gets discounted by the same amount. We use one of the key assumptions from Kahneman and Tversky (1979) to weight the externality losses (in case a negative externality is selected) simply, making it count twice as much as the gains.

The ABRL framework we implemented employs a decision rule grounded in the concept of attractions, which are essentially weights assigned to strategies. These attraction values are updated based on the payoffs from successful computer selections (C. Camerer & Hua Ho, 1999; Erev & Roth, 1998). An extension of the ABRL agent, the ϵ-greedy strategy, emerges from the multi-armed bandit framework (Robbins, 1952; Robbins & Monro, 1951). It introduces an exploration-exploitation trade-off by incorporating an exploration rate, ϵ. This rate represents the likelihood of the agent opting for a random strategy. As the experiment unfolds, ϵ decays from its initial value to a minimum threshold, emphasizing exploitation in its latter stages. In contrast, the UCB method also navigates the exploration-exploitation dilemma by factoring in the current reward estimates and their uncertainties. By integrating a confidence interval with each action's value, UCB balances exploring less familiar actions with exploiting known, high-reward actions, thereby optimizing cumulative gains (Auer, Cesa-Bianchi, & Fischer, 2002; Sutton & Barto, 2018). For a deeper dive into the theoretical underpinnings of each model, refer to Appendix 6.5; for a detailed technical breakdown, see Appendix 6.5.1.

In the simulation model, the three types of agents—ABRL, ϵ-greedy, and UCB—are all tasked with the same objective, mirroring the original design: they sample for information and then select a computer based on a Bayesian update rule. Instead of directly learning how to choose the best computer (since its assignment is randomized each round), the agents learn the sampling process to ensure that the subsequent computer selection yields maximized rewards. We added a fourth agent that always draws random samples for comparison purposes. The feedback part is not considered in the simulation, as it is necessary to provide feedback for the RL updating rules in any case.

6.3 Experimental Results

This section presents the results of the human subject experiment. First, we see the results for total sampling S per round across the treatment dimensions and how these results are driven by gender. We continue with the analysis of motivated sampling. First, we analyze the extensive margin of unequal sampling between the options A and B. Then, we split it by whether or not the externality option A_{ext} was selected. We then analyze the intensive margin of this unequal sampling, split by whether or not the externality option A_{ext} was selected. Here, we measure the intensive margin of motivated sampling $MotSamp$. Furthermore, we analyze the decision behavior and scoring success of subjects. We measure the time subjects actually take for sampling. Lastly, we show the accuracy of the stated posterior belief.

6.3.1 Total Sampling per Round

First, we assess the total sampling behavior S per round. Here, we will look at the effect of the externalities and feedback treatments on the total number of samples a subject created for both computers in one round.

We find that participants sample most in the negative externality (12.18 samples per round) and equally in the no- and positive externality treatments (11.01 and 10.95). Across the feedback dimension, the participants sample more, the more detailed feedback they receive: Bayes feedback (12.00) > outcome feedback (11.54) > no feedback (10.61).

The plots in figure 6.3 show that these effects are driven by gender. Female participants sampled significantly more than male participants in the negative externality and Bayes feedback treatments. This insight shows that women make higher sampling efforts when the context is most salient, either through a negative externality or detailed feedback on their stated posterior belief.

6.3.2 Motivated Sampling

After this effect of total sampling per round, we turn to motivated sampling within a round. We define motivated sampling as the tendency to sample additionally *because* of subjective preferences. To identify this behavior, we need to perform an analysis in multiple steps.

First, we will identify the extensive margin of whether subjects sampled unequally for the two available options. Then, we will split this by whether the externality option (A_{ext}) was selected or the non-externality option (B_{non}). This split will give us a score of the

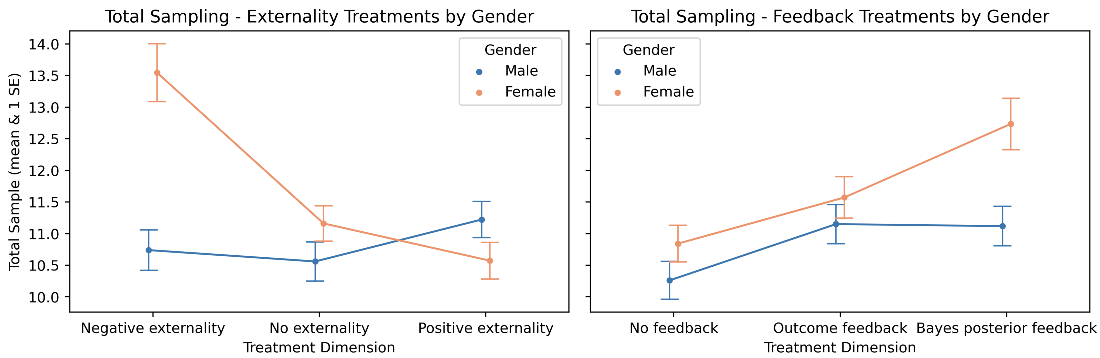


Figure 6.3: sampling by gender - externality and feedback treatments

extensive margin of motivated sampling, but more importantly, we use it to move to the intensive margin of how many subjects sample more when they sample unequally. We will then also split by whether the externality option A_{ext} was selected or not, which gives the crucial comparison of the sampling behavior when subjects deem A_{ext} correct compared to when they deem B_{non} correct. This allows us to identify the additional sampling from an option out of a subjective preference for doing so. Out of this, we will calculate this score of motivated sampling in a round $MotSamp$, split by type of externality and gender.

6.3.3 Extensive Margin Unequal Samples

Many subjects might use a strategy to constantly sample equally from options A and B. Those subjects will not show a behavior of motivated sampling within a round. Nevertheless, many subjects might sample unequally from the two options, resulting in unequal sample sizes $s_A \neq s_B$. We will first turn to the extensive margin of this unequal sampling behavior, whether a subject did sample unequally or not.

We measure for each participant how many of the 20 rounds they show an unequal sample strategy $s_A \neq s_B$. Figure 6.4 plots this fraction of rounds in which a subject showed an unequal sample. 23% subjects show a fraction of 0, so never sampling unequally, i.e., constantly sample the same number of information from both sources. In total, 49.7% of all subjects show unequal samples in only a quarter of the rounds or less. This effect is drastically more than compared to 6.8%, who constantly sample unequally, and 25.6%, who show unequal samples in at least three-quarters of all rounds. This pattern is stable over all treatments.

6.3.4 Extensive Margin Unequal Samples with Externality Selection

Table 6.2 shows the extensive margin of how many observations in a treatment exhibit an unequal sample. In the first column, overall by treatment, where treatments with an externality show more unequal samples and especially the negative externality treatment shows statistically significant more unequal sampling than with positive externality ($p < 0.01$) and without externality ($p < 0.001$) using a Chi-Squared Test. Columns two and three split this overall value in whether the externality or non-externality option was selected. Now, we can calculate the extensive margin of motivated sampling.

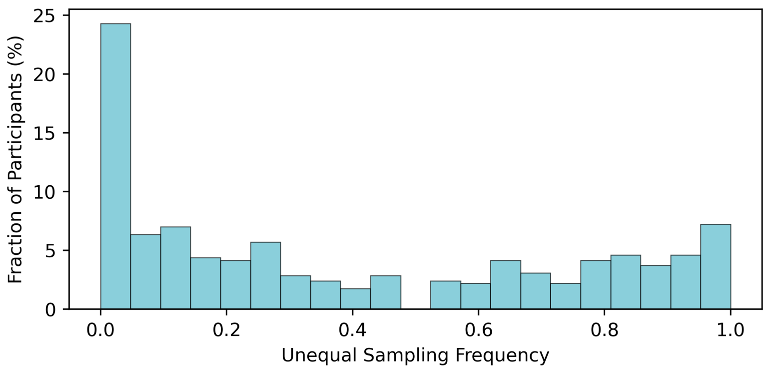


Figure 6.4: Unequal sampling by subjects

With a negative externality, we see $42.7 - 40.3 = 2.4$ percentage points more unequal samples when the externality is selected. With a positive externality, we observe $38.3 - 37.9 = 0.4$ percentage points more unequal samples, so in both cases, a relatively mild extensive margin of motivated sampling.

Externality	Overall	Ext select	Non ext select
No	37.1%	-	-
Negative	41.5%	42.7%	40.3%
Positive	38.1%	38.3%	37.9%

Table 6.2: Extensive Margin Unequal Sampling

6.3.5 Intensive Margin of Motivated Sampling

Now, we turn to the intensive margin, so how many more subjects sample for one of the options when they showed a $s_A \neq s_B$ unequal sampling. For this, we calculate the Δ-sample, Δs, which is the difference between the sample for the selected option, s_{select}, and the option that was not selected, $s_{NonSelect}$, $\Delta s = s_{select} - s_{NonSelect}$.

Table 6.3 shows Δs in the first column by treatment. Here, we see motivated sampling as the intensive margin with no externality is with $\Delta s = 0.419$ sample on average lower than both intensive margins with negative (0.876) and positive (0.912) externality.

Columns two and three split Δs by whether the option with externality was selected, Δs_{ext}, or without, Δs_{non}.

Externality	Δs	Δs_{ext}	Δs_{non}
No	0.419	-	-
Negative	0.876	0.842	0.91
Positive	0.912	1.769	-0.075

Table 6.3: Intensive Margin Unequal Sampling Δs

With Δs_{ext} and Δs_{non}, we can now calculate the intensive margin of motivated sampling as the additional sampling when the externality option A_{ext} was selected as $MotSamp = \Delta s_{ext} - \Delta s_{non}$. With a negative externality, we see almost no motivated sampling within a round ($MotSamp = 0.842 - 0.91 = -0.068$) as the margin of unequal sampling is relatively similar regardless of whether subjects select the externality.

However, with a positive externality, there is a striking difference. Here, we see a robust case of motivated sampling of $MotSamp = 1.769 - (-0.075) = 1.844$. This large effect shows that subjects sample much more additionally from the positive externality option, hence "like" sampling from it when they deem it correct.

Table 6.4 splits this analysis of motivated sampling by gender. With both types of externalities, we observe that male participants show a stronger liking for the "nicer" option than female participants, i.e., the option with positive externality or the option *without* negative externality, respectively.

Externality	Overall	Male	Female
Negative	-0.068	-0.201	0.200
Positive	1.844	1.925	1.637

Table 6.4: Motivated sampling $MotSamp$

6.3.6 Decision and Scoring behavior

Next to the sampling behavior, we are also interested in how externalities and feedback influence the decision and scoring behavior. Here, we should note that the correct option was always determined randomly, so the correct baseline is always a 50/50 split. Moreover, the option subjects select is relatively equally balanced between the options with and without externality with a negative externality (49.0% to 51.0%). With a positive externality, subjects decide in 53.3% for the externality option. This is significantly more than the balanced split ($p < 0.001$) in a binomial test, as seen in table 6.5, mirroring the motivated sampling into decision behavior.

Externality	Externality Select	Non-ext select
Negative	49.0%	51.0%
Positive	53.3%	46.7%

Table 6.5: Select Ext / Non-Ext per Externality Treatment Dimension

Interestingly, this does not translate into meaningful differences in correct decisions, as all treatments hover around 75% accuracy in their decisions, see table 6.6. This effect is also very stable throughout the experiment of 20 rounds.

Treatment	Scoring
Overall	75.2%
Feedback	
Outcome	74.5%
Bayes	77.0%
No	74.2%
Externality	
Negative	75.3%
Positive	74.8%
No	75.5%

Table 6.6: Scoring across treatments

6.3.7 Page Time Analysis

Furthermore, as an exploratory analysis, we investigate the time subjects take to complete the core sampling task. Figure 6.5 reports on the x-axis the seconds subjects spend on the page for the main task from the start of sampling till confirmation of their decision. We plot the 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile and the mean of the seconds they spent on the page for the task. An interesting finding for further research is that even though subjects in the negative externality treatments sample more, they spend less time on the task than subjects in the other treatments. Potential explanations could be that they take less time to evaluate the samples through more sampling or that the unpleasant presence of the negative externality prompts subjects to be faster in their evaluation process, which would show another form of motivated effort distribution, a hypothesis for further research.

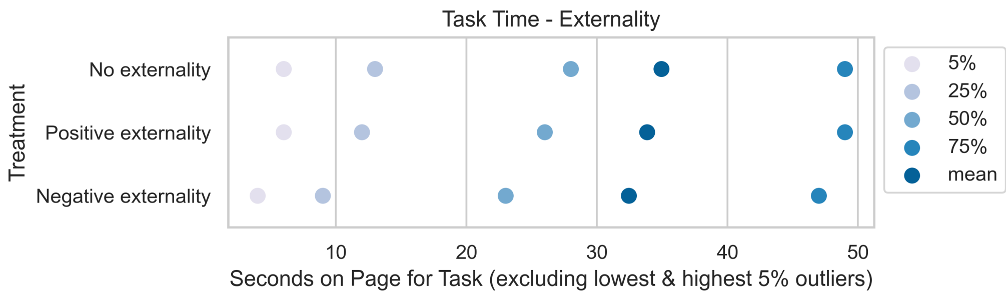


Figure 6.5: Task time

6.3.8 Posterior Beliefs

After subjects completed the task as part B of a round, we asked them, "Please let us know how likely it seems to you that your choice was correct." Table 6.7 reports the percentage point difference between their stated belief and the rational Bayesian posterior. We see that subjects, on average, gave a lower estimation than the rational Bayesian posterior, where the subjects who receive the Bayes feedback are closest to the rational posterior since they get feedback on their stated posterior and the Bayesian posterior in every round and can learn from it. Interestingly, subjects with only outcome feedback significantly differ significantly from those without feedback. As pre-registered, we excluded participants who stated a posterior belief lower than 50% as a potential sign they did not seriously think about the question.

Treatment	Difference
Overall	-4.4%
Feedback	
Outcome	-6.7%
No	-4.2%
Bayes	-2.3%

Table 6.7: Difference Stated- to Bayesian-Posterior

6.3.9 Agent-Based Simulation Insights

In this chapter, we examine the outcomes of our task simulation using different variations of RL agents. In Table 6.8, we have consolidated the results from the simulation model across all agent classes. The rows labeled ”Experiment” represent metrics corresponding to human participants. Notably, the ”reward” metric for human participants was computed using the identical formula employed for the agents, ensuring a consistent basis for comparison. This calculated ”reward” should be differentiated from the actual rewards that participants received, which may have been in the form of payoffs or monetary compensations. Additionally, we introduced a ”Random” agent that samples randomly within the pre-defined limits to serve as a benchmark. Significantly, all the RL models and the human participants consistently outperformed this random benchmark, underscoring the efficacy of the RL approach in the task.

Method	Externality	Reward	Samples	Posterior	Correct	Freq. Externalities
Experiment	negative	0.38	12.17	0.74	0.75	0.12
Experiment	neutral	0.48	11.06	0.74	0.75	0.00
Experiment	positive	0.56	10.99	0.73	0.74	0.40
ABRL	negative	0.37	18.26	0.77	0.78	0.11
ABRL	neutral	0.47	13.23	0.73	0.72	0.00
ABRL	positive	0.56	12.79	0.72	0.72	0.34
ABRL ϵ-Greedy	negative	0.32	25.15	0.83	0.84	0.08
ABRL ϵ-Greedy	neutral	0.38	23.97	0.83	0.84	0.00
ABRL ϵ-Greedy	positive	0.48	21.65	0.80	0.80	0.40
UCB	negative	0.39	15.04	0.73	0.74	0.14
UCB	neutral	0.47	12.27	0.71	0.70	0.00
UCB	positive	0.55	16.13	0.74	0.76	0.37
Random	negative	0.34	24.04	0.83	0.83	0.08
Random	neutral	0.38	23.66	0.83	0.83	0.00
Random	positive	0.49	23.88	0.85	0.84	0.42

Table 6.8: RL simulation model - Results consolidated by average

Several salient observations arise in our comparative study between human participants and a range of reinforcement learning strategies. When juxtaposed with attraction-based reinforcement learning (ABRL) agents, ABRL with ϵ-greedy exploration, and the Upper Confidence Bound (UCB) method, human participants tended to sample less frequently. Despite this reduced sampling, their performance remained on par with the RL agents, suggesting a heightened efficiency in their decision-making process. This insight is further illustrated in figure 6.6, showcasing the distributions of samples and cumulative rewards for each method.

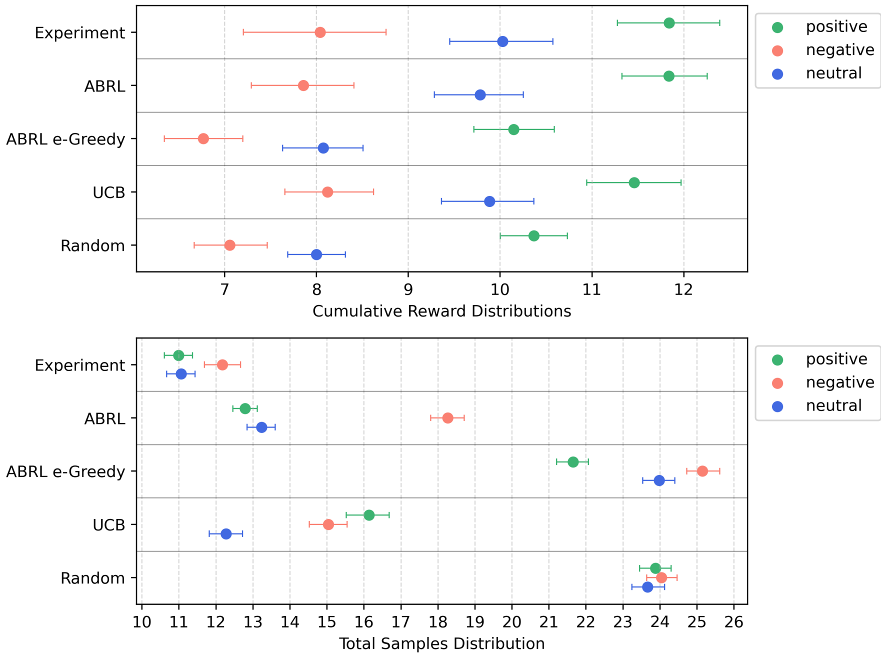


Figure 6.6: Distributions of Samples and Cumulative Rewards

The ABRL strategy, which can be viewed as a more optimization-centric method, showcased a performance profile comparable to that of the human participants. Both in terms of sampling frequency and rewards, fact further illustrated in figure 6.7, comparing average samples and cumulative rewards. In addition, externalities played a significant role in influencing sampling behavior. Introducing a negative externality exerted a more pronounced impact on sampling frequency than its positive counterpart. This trend was consistently observed in both human participants and the simulation, indicating a universal aversion to potential losses or penalties.

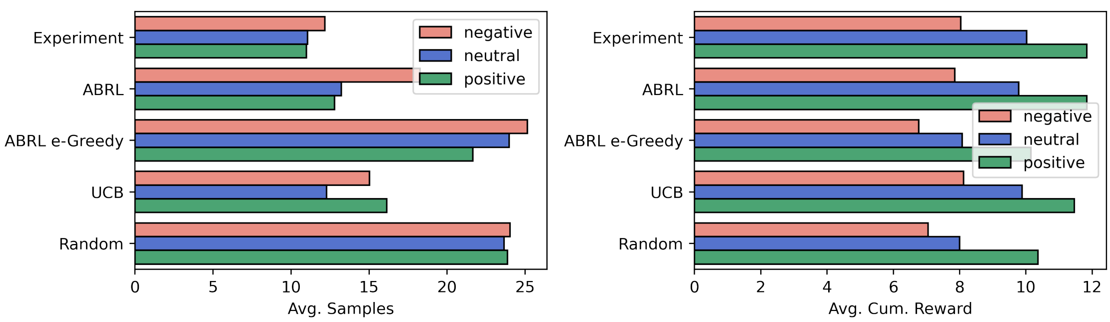


Figure 6.7: Comparison of Mean Samples and Cumulative Rewards

We also noted a nuanced relationship between sampling, certainty, and rewards. While increased sampling enhances the posterior certainty, it comes at a cost. Given the associated costs, excessive sampling can erode the net rewards, highlighting the balance decision-makers must strike between gathering information and acting efficiently.

In summation, the simulation provides an additional layer of understanding that complements the primary analysis. It consistently demonstrates the reinforcement learning behavior in human participants, emphasizing their inclination to act optimally. This exploration further delineates the nuanced relationship between data sampling, externalities, and decision-making, highlighting parallels between human intuition and algorithmic strategies and enriching our comprehension of human-like decision processes.

6.4 Discussion

In this study, we have investigated the intersection of information sampling and motivated reasoning, which we term motivated sampling.

We find substantial evidence for motivated sampling, which emerges in different forms. First, female participants sample significantly more in context-rich environments, especially when the externality is negative. Furthermore, subjects show a strong sense of motivated sampling when they deem the option with a positive subjective preference correct. Here, they also show a behavior of "liking" to sample from it. This behavior of motivated sampling is more strongly pronounced for male than female subjects. This effect translates into a tendency to select the option associated with the positive externality more often than it would be objectively correct. As complementary findings, we observe that subjects use the least time for the task in the negative externality treatment, even though they sample the most. If subjects receive feedback on the rational Bayesian posterior, their stated posterior belief is very close to it. When subjects receive feedback on the outcome of their decision, the stated posterior belief is more distant from the Bayesian posterior than when subjects receive no feedback.

By employing a simulation model, we gained further insights into the dynamics of motivated sampling. The model served as a valuable tool to validate and extend our empirical findings, revealing consistent reinforcement learning behaviors in participants as they sought to act optimally. The simulation underscored the intricate balance participants strike between data sampling, externalities, and decision-making. It also highlighted the parallels between human intuition and algorithmic strategies, adding depth to our understanding. This computational approach complements our primary analysis and offers a framework for future studies to explore the nuances of motivated sampling in various contexts.

With this study and its findings, we contribute to a better understanding of information sampling behavior as a specific application of the broader field of motivated reasoning. This contribution indicates different cost functions of sampling depending on externality, relating to Petitet et al. (2021) and Kelly et al. (2021).

Our study shows that the amount and direction of effort spent on information sampling in a decision-making context is greatly influenced by subjective criteria. The gender-based differences, particularly among female participants in specific treatments, indicate that context might influence information acquisition differently across genders, potentially having implications in areas like marketing, education, or policy-making.

The study naturally carries several limitations. The study's online nature might introduce biases, as participants could be influenced by external factors while doing the experiment not present in controlled lab settings. The list of organizations provided as externality is inherently incomplete to elicit strong subjective preferences for all subjects. It is expected that some participants had no "strong feelings" about any of them, reducing the studied effect of externalities. Furthermore, experienced subjects might not have responded truthfully to the questions about the organizations, anticipating that they might play a role later and therefore just indicating organizations they feel indifferent about. This would imply that our findings provide a lower bound of the effect of externalities in motivated sampling.

Future studies could explore the underlying psychological factors driving the observed behaviors, especially in negative externality scenarios. It would also be intriguing to analyze further the gender differences and what might drive them. Additionally, expanding the study to diverse demographic groups or introducing more complex decision-making tasks could provide richer insights. The behavior termed motivated sampling can be seen as a fundamental, underlying mechanism of confirmation bias, which future studies should explore specifically.

This research highlights the role of objective and subjective criteria in information sampling. The findings, particularly regarding gender differences and the influence of personal values, underscore the importance of the phenomenon of motivated sampling.

6.5 Appendix

Methods and Formulas

This chapter provides the mathematical formalizations for the methods used in calculating Bayesian posteriors and the RL-based decision rules for the agent-based models.

Bayesian Posterior

Our design introduces two main assumptions. Firstly, we assume that each computer generates a number from one of two distinct distributions: high or low. Secondly, before any data sampling, both computers are presumed to have an equal probability of drawing from either the high or low distribution, with the prior probability set at an even 50/50. In our notation, HL indicates that computer 1 is sampling from the high distribution while computer 2 is drawing from the low distribution. Given an array X representing all sampled numbers from both computers, we can apply Bayes' theorem to determine $P(HL \mid X)$, which represents the probability that computer 1 is from the high distribution and computer 2 is from the low distribution:

$$P(HL \mid X) = \frac{P(X \mid HL)P(HL)}{P(X \mid HL)P(HL) + P(X \mid LH)P(LH)} \tag{6.1}$$

This formula was used to update the prior beliefs in the experiment and in the simulation model.

Reinforcement Learning Agents

The ABRL agent operates within the Attraction-Based Reinforcement Learning framework, leveraging the Softmax algorithm to determine the optimal sample size. Two pivotal components underpin this methodology. Initially, we have the attraction update given by:

$$A_{s_i}(n) = \phi A_{s_i}(n-1) + I(s(n) = s_i)\pi_{s_i}(n) \tag{6.2}$$

Subsequently, the decision-making process employs the Softmax criterion, expressed as:

$$P_{s_i}(n+1) = \frac{e^{\lambda A_{s_i}(n)}}{\sum_{k=1}^{m} e^{\lambda A_{s_k}(n)}} \tag{6.3}$$

Here, s denotes the sample size selection action, n represents the period, and π is the payoff associated with a specific computer selection action. The model incorporates two parameters: ϕ (the Recency parameter), which acts as a discount rate for attractions, emphasizing the significance of previous experiences, and λ, the Softmax update weight. The latter parameter, λ, gauges the sensitivity or weight of attractions. A value of $\lambda = 0$ implies that attractions hold no relevance, while a high λ value accentuates their importance.

Building on this, the ϵ-greedy strategy introduces a nuanced balance between exploration and exploitation. Here, exploration corresponds to a random choice, while exploitation adheres to the Softmax rule, as delineated in the ABRL:

$$s_i(n) = \begin{cases} \text{random action} & \text{with probability } \epsilon \\ \text{softmax rule} & \text{with probability } 1-\epsilon \end{cases} \tag{6.4}$$

The UCB (Upper Confidence Bound) model integrates three pivotal components to guide its decision-making process. Initially, the model employs an action selection rule represented by

$$s_i(n) = \arg\max_s \left[Q(s) + \sqrt{\frac{2\log N}{n(s)}} \right] \tag{6.5}$$

This rule is designed to determine the optimal action by maximizing the estimated value of an action, $Q(s)$, while accounting for an uncertainty term. This ensures a balance between exploring actions with less certainty and exploiting those with higher known rewards.

Subsequently, the model updates its beliefs using the rule:

$$Q_{n+1}(s) = Q_n(s) + \frac{1}{n(s)}\left[R_n(s) - Q_n(s)\right] \tag{6.6}$$

This equation refines the estimated value of an action based on the latest rewards received, ensuring that the model's beliefs are always in line with the most recent observations.

Lastly, the model generates the upper confidence bounds, denoted as UCB, using the equation:

$$UCB(s) = Q(s) + \sqrt{\frac{2 \log N}{n(s)}} \tag{6.7}$$

This calculation provides a measure of the potential upside of each action, taking into account both its estimated value and the uncertainty surrounding it. This ensures that the model remains explorative, especially in the face of uncertainty, while also being opportunistic when the rewards are more predictable.

6.5.1 Technical Remarks

The experiment was programmed using oTree (D. L. Chen et al., 2016). The simulation models were programmed in Python and employed the following base parameter described in table 6.9.

Parameter	Value	Description
Simulated agents	100	Number of agents simulated per externality type
Periods	21	Same amount of periods as the experimental task
Maximum samples	25	Maximum samples allowed
Base payoff	1	Amount of points per correct answer
Sample costs	0.019	The amount a sampling action costs, relative to a point (base payoff)
Externality value	0.25	The value of an externality realization E, relative to a point

Table 6.9: Parameters for the Agent-Based model

In addition, the ABRL parameters ϕ and λ were tuned using the Optuna library for hyperparameter optimization (Akiba et al., 2019), following a ”maximize reward” target. The ϵ-greedy variant of the ABRL introduces an additional ϵ parameter that unfolds into three more parameters when using dynamic values, decaying ϵ over time. The three additional parameters are then ϵ_{start}, which is the upper bound value, ϵ_{end}, the lower bound value, and ϵ_{decay}, which governs the rate at which ϵ_{start} decays to ϵ_{end}. These parameter values were also tuned for each case using Optuna. In every case, The optimization process employed 450 trials for each variant.

6.5.2 Design Details

This section contains the experiment screens shown to the participants during the tasks. Figure 6.8 and 6.9 show the unbiased prior estimation task before and after clicking. Figure

Figure 6.10 shows the posterior estimation task, also with an unbiased slider.

For the externalities design, figure 6.11 contains the ”no externality variant”, while figures 6.12 and 6.13 contain the positive and negative externality design examples, respectively.

In the case of negative externalities, we gave participants in this treatment the chance to opt out of the experiment since it could involve the action of donating money to a disliked

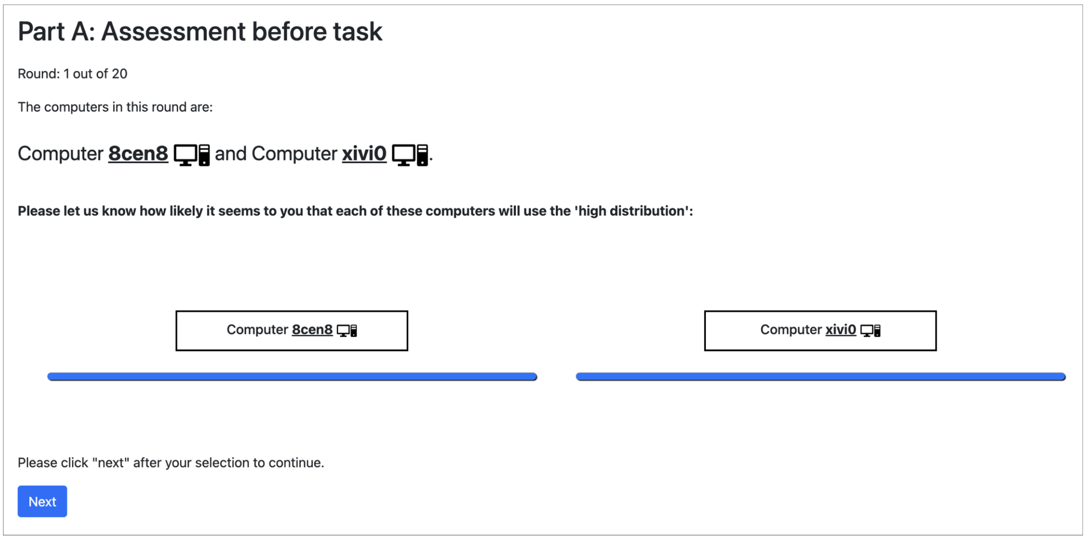


Figure 6.8: Prior before click (no-default slider)

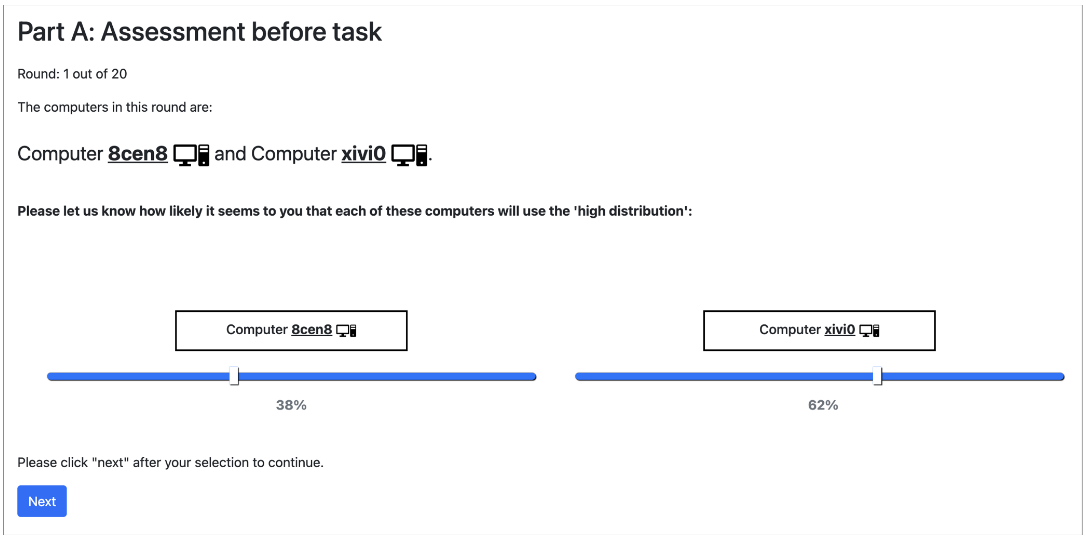


Figure 6.9: Prior after click

organization. The opt-out screen is shown in figure 6.13. Participants who opted out of the experiment were still paid the show-up fee.

Lastly, figure 6.15 contains examples of the different variations of feedback provided. From left to right, we have ”no feedback,” ”outcome feedback,” and ”Bayesian posterior feedback.”

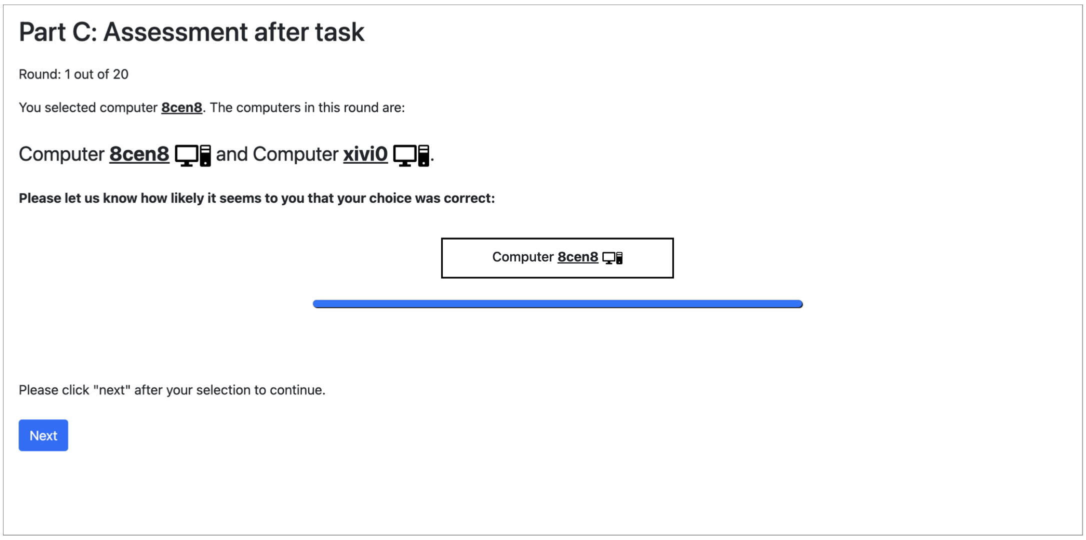


Figure 6.10: Posterior

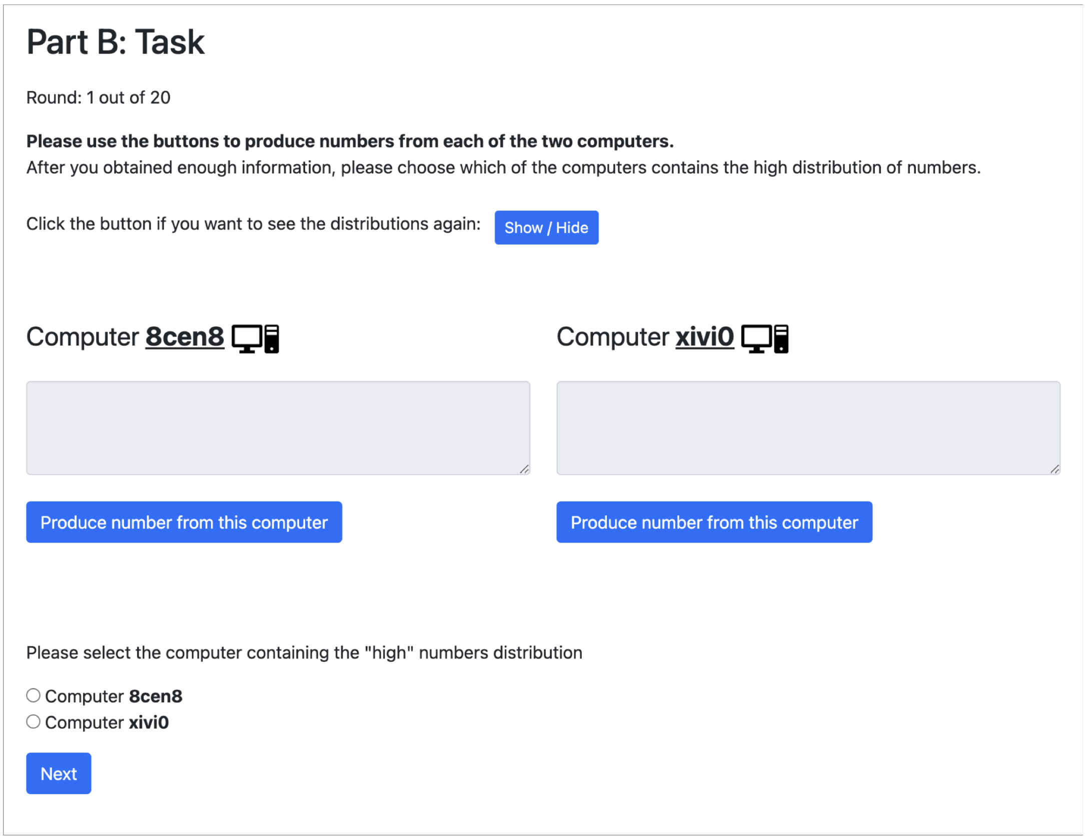


Figure 6.11: No externalities

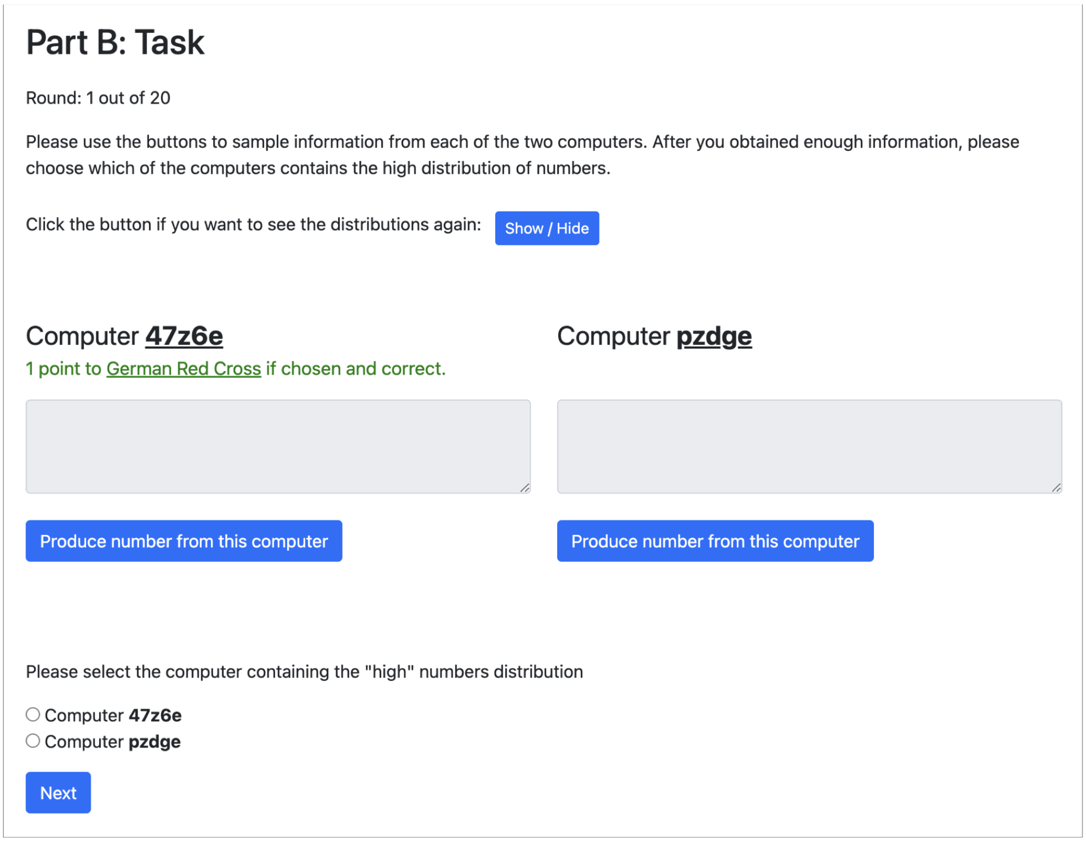


Figure 6.12: Positive externalities

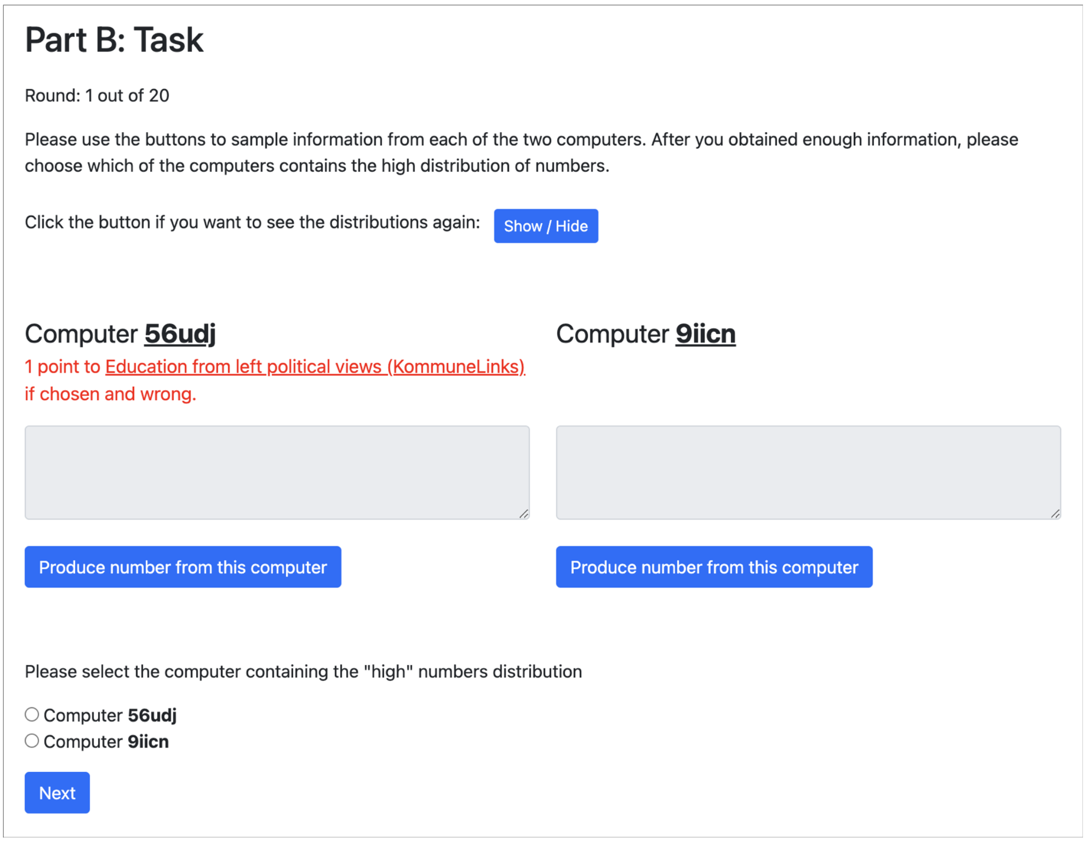


Figure 6.13: Negative externalities

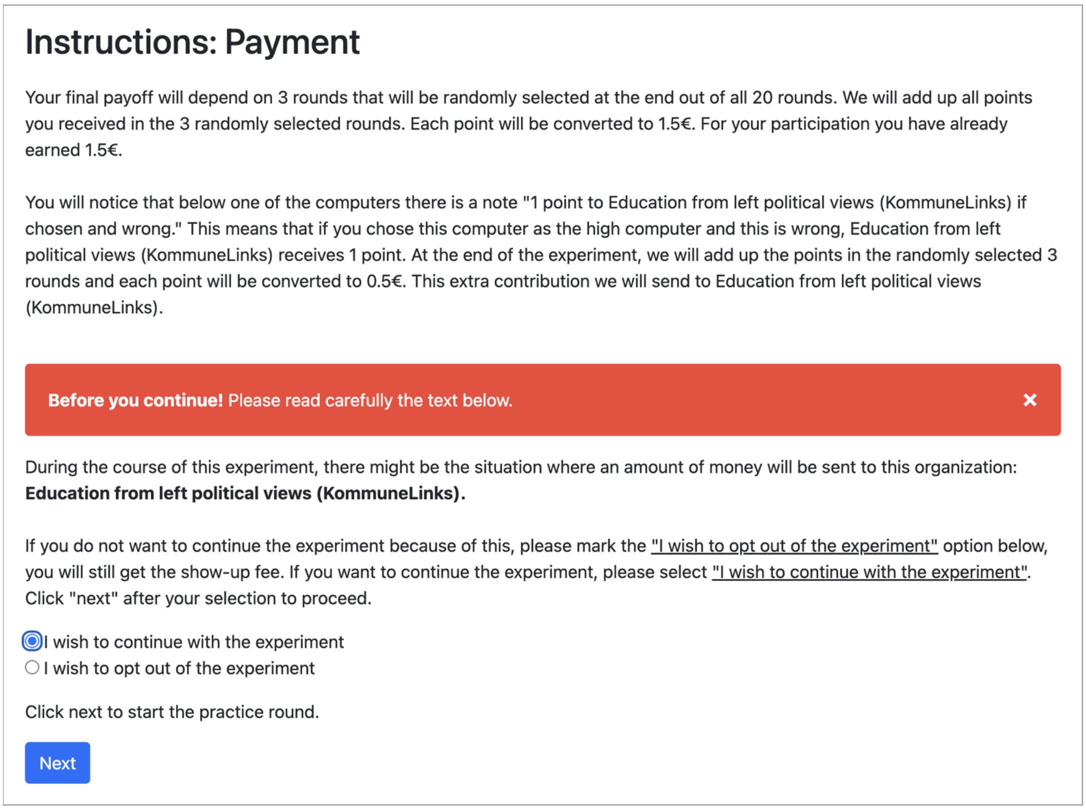


Figure 6.14: Opt-out in case of negative externalities

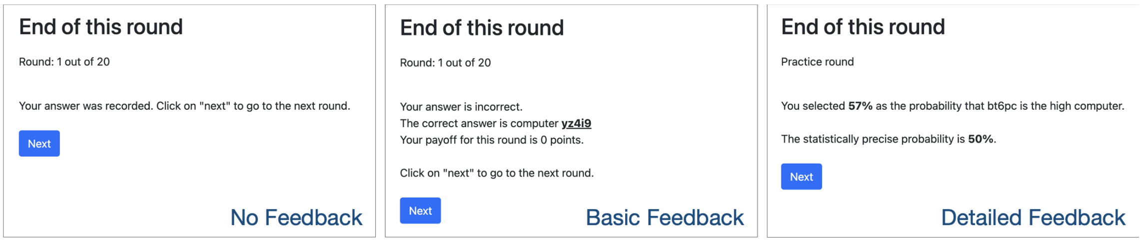


Figure 6.15: Feedback

7. Conclusions

The body of work consolidated in this research project aimed at advancing the field of economics through behavioral-focused analyses, experimentally validated insights, and advanced methods that enable deep exploration of decision-making phenomena. I trust that the findings from the five introduced papers can guide practitioners and chart new research directions. From the experiment designs to the algorithms developed, an emphasis was placed on transparency, robustness, generalizability, and reproducibility. A concerted effort was made to provide clarity in the assembly of complex methodological constructs by drawing insights from other analytical fields, such as computer science and causal inference.

In the journey for new findings and insights, this research has been guided by a commitment to understanding the psychological elements of human decisions. I have explored how individuals navigate economic scenarios, considering a wide array of factors, from personal traits to contextual elements and strategic interaction. The experimental designs in laboratory and online environments have been tailored to capture authentic behavioral responses, generating results that enrich the understanding of economic decisions.

My intended contributions to the field aim to help bridge the gap between traditional economic theories and observed behavioral outcomes through contemporary designs, actionable findings, and methodological innovation. With objectives focused on integrating behavioral, experimental, and computational economics to enhance understanding of human decision-making, the essence of my envisioned contributions is outlined in the following points:

1. A novel, open-source simulation model using genetic algorithms to simulate evolving strategic scenarios. This includes several different game types, behavioral profiles, and decision rules.

2. Behavioral modifications to stationary equilibrium concepts based on loss aversion and fairness, which substantially improve their predictive performance over the seminal literature.

3. A novel analysis of dark personality traits and framing in the context of inspection games, with an extensive analytical approach involving advanced machine learning methods and causal inference.

4. A comprehensive study of algorithm aversion with novel treatment design conditions, such as payment, automation, and method description, delivering robust and significant insights on human-computer interaction using advanced machine learning techniques.

5. A new perspective on motivated sampling, revealing gender-based distinctions in information acquisition within decision-making, and utilizing a reinforcement learning-based simulation model to validate and deepen our comprehension of these behaviors in varying conditions.

From a theoretical perspective, these contributions include introducing innovative models that enrich existing economic theories by offering novel perspectives on strategic behavior, equilibrium dynamics, and the role of behavioral traits in economic interactions. Concerning practical contributions, the presented findings are intended to provide real-world relevance and offer actionable insights for practitioners, policymakers, and decision-makers. Examples include guidance on designing systems and interventions that effectively deal with aversion, cheating, and asymmetric information while accounting for pre-defined personality traits and different behavioral profiles. In essence, these implications collectively contribute to enhancing decision-making processes and outcomes in various economic contexts.

The relevance for practitioners lies in the tangible applications of the research findings and methods. The open-source simulation model using genetic algorithms offers a predictive tool for anticipating strategic interaction dynamics and competitive behavior. Insights into behavioral modifications to equilibrium concepts enable the crafting of effective incentive structures and policies that resonate with human behavioral traits. The exploration of dark personality traits in inspection games can guide organizations in designing robust tools and policies where dishonesty-related behavior is relevant, especially in trust-critical subjects, for example, fraud detection. The study on algorithm aversion provides technology companies with strategies to enhance user trust in AI-driven tools, ensuring smoother adoption and optimal behavior and considering specific personality profile types. Lastly, understanding motivated sampling can empower organizations to refine their information presentation strategies, optimizing their messaging to align with consumer decision-making processes with different characteristics.

From a holistic point of view, there are limitations that I recognize and aim to address. The game-theoretical experiments primarily focus on specific game types, potentially limiting the applicability of findings to broader decision-making contexts. In addition, the reliance on simplified behavioral models and experimental designs may not capture the full complexity of human choices. It is also noteworthy that my samples of participants, which consisted mainly of students, may only partially represent diverse populations. Lastly, some methods employed, especially regarding machine learning models, can be computationally demanding, potentially limiting real-time applicability. Recognizing these limitations underscores the need for future research to address these constraints and expand the breadth and depth of insights in the field.

Embarking on this research journey has been both enlightening and challenging, an unforgettable voyage through scientific and personal discovery. From the first steps of formulating research questions and analyzing the current literature, combined with the expert collaboration of the coauthors, as mentioned earlier, this process offered valuable insights into the gaps and opportunities that shaped my direction. My approach's multidisciplinary nature introduced new perspectives on analyzing and documenting findings from economic experiments. The path taken from conception until this moment is marked with personal growth, beautiful experiences, and a drive to contribute to the higher purpose of science.

Future research could consider testing the introduced models and methods using different designs and participant pools in varied scenarios. The novel simulation and equilibrium models can be extended to incorporate further components representing other facets of human behavior, strengthening the connection between theory and practice. Moreover, approaching different problems using similar methodologies, such as machine learning models and advanced causal inference techniques, can validate and sharpen the usability of such techniques to analyze complex experimental data. I employed generalized scenarios in the two more contextual experiments of algorithm aversion and motivated sampling. Such experiments could be extended to represent specific situations and test different context frames that reveal other sides of the underlying phenomena.

To conclude, throughout this research, I have systematically investigated the intersections of economics, behavioral analysis, and computational methods. Across five independent but connected papers, I have presented evidence and analyses that aim to expand our understanding of economic decision-making processes. My methodological approach focused on painting a comprehensive picture by combining traditional analytical frameworks with cutting-edge methods in machine learning, agent-based simulations, and causal inference. As with any scientific endeavor, these findings also open avenues for further questions and research. I hope this body of work serves as both a valuable contribution to current literature and a foundation for future investigations. It is my intent that this research provides a clearer, more rigorous understanding of how individuals make economic decisions in varied contexts.

8. General Appendix and Additional Information

8.1 Conference Presentations

Table 8.1 displays the list of conference and workshop paper presentations produced and executed in the context of this thesis.

Paper	Conference	Date	Place
Simulating Economic Learning in Dynamic Strategic Scenarios with a Genetic Algorithm	12th International Conference of the French Association of Experimental Economics (ASFEE)	Jun 2022	Lyon, France
Stationary Equilibria in Behavioral Game Theory: An Experimental Analysis of Inspection Games	12th International Conference of the French Association of Experimental Economics (ASFEE)	Jun 2022	Lyon, France
Stationary Equilibria in Behavioral Game Theory: An Experimental Analysis of Inspection Games	The European Economic Science Association (ESA) Conference	Aug 2022	Bologna, Italy
Simulating Economic Learning in Dynamic Strategic Scenarios with a Genetic Algorithm	57th Hohenheimer Oberseminar (HOS)	Dec 2022	Kleve, Germany
Stationary Equilibria in Behavioral Game Theory: An Experimental Analysis of Inspection Games	57th Hohenheimer Oberseminar (HOS)	Dec 2022	Kleve, Germany
Trust in the Machine: How Contextual Factors and Personality Traits Shape Algorithm Aversion and Collaboration	BSE Computational and Experimental Economics Workshop	Jun 2023	Barcelona, Spain
Motivated Sampling of Information: Analysis With Experimental Data and Agent-Based Modeling Within a Bayesian Framework	BSE Computational and Experimental Economics Workshop	Jun 2023	Barcelona, Spain
Motivated Sampling of Information: Analysis With Experimental Data and Agent-Based Modeling Within a Bayesian Framework	Subjective Probability, Utility and Decision Making (SPUDM)	Aug 2023	Vienna, Austria

Table 8.1: List of Paper Presentations in Conferences and Workshops

8.2 Supplementary Files

The following pages include the additional material to be appended to this thesis. In the first supplementary sheet, information about my contributions to each paper is provided. Next, an affidavit and declaration of originality is presented.

References

Abbink, K., & Hennig-Schmidt, H. (2006). Neutral versus loaded instructions in a bribery experiment. *Experimental Economics*, *9*, 103–121.

Abbink, K., Irlenbusch, B., & Renner, E. (2002). An experimental bribery game. *Journal of Law, economics, and organization*, *18*(2), 428–454.

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining* (pp. 2623–2631).

Amit, Y., & Geman, D. (1997). Shape quantization and recognition with randomized trees. *Neural computation*, *9*(7), 1545–1588.

Andreoni, J. (1995). Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. *The Quarterly Journal of Economics*, *110*(1), 1–21.

Andreozzi, L. (2004). Rewarding policemen increases crime. another surprising result from the inspection game. *Public choice*, *121*, 69–82.

Arifovic, J., & Ledyard, J. (2004). Scaling up learning models in public good games. *Journal of Public Economic Theory*, *6*(2), 203–238.

Arifovic, J., & Ledyard, J. (2011). A behavioral model for mechanism design: Individual evolutionary learning. *Journal of Economic Behavior & Organization*, *78*(3), 374–395.

Arifovic, J., & Ledyard, J. (2012). Individual evolutionary learning, other-regarding preferences, and the voluntary contributions mechanism. *Journal of Public Economics*, *96*(9-10), 808–823.

Athey, S. (2018). The impact of machine learning on economics. In *The economics of artificial intelligence: An agenda* (pp. 507–547). University of Chicago Press.

Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, *113*(27), 7353–7360.

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*, 685–725.

Athey, S., Tibshirani, J., & Wager, S. (2018). *Generalized random forests.*

Athey, S., & Wager, S. (2019). Estimating treatment effects with causal forests: An application. *Observational studies*, *5*(2), 37–51.

Auer, P., Cesa-Bianchi, N., & Fischer, P. (2002). Finite-time analysis of the multiarmed bandit problem. *Machine learning*, *47*, 235–256.

Avenhaus, R. (2004). Applications of inspection games. *Mathematical Modelling and Analysis*, *9*(3), 179–192.

Avenhaus, R., Canty, M., Kilgour, D. M., Von Stengel, B., & Zamir, S. (1996). Inspection games in arms control. *European Journal of Operational Research*, *90*(3), 383–394.

Avenhaus, R., Von Stengel, B., & Zamir, S. (2002). Inspection games. *Handbook of game theory with economic applications*, *3*, 1947–1987.

Axelrod, R. (1997). *The complexity of cooperation: Agent-based models of competition and collaboration* (Vol. 3). Princeton University Press.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *science*, *211*(4489), 1390–1396.

Axelrod, R., et al. (1987). The evolution of strategies in the iterated prisoner's dilemma. *The dynamics of norms*, 1–16.

Azucar, D., Marengo, D., & Settanni, M. (2018). Predicting the big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and individual differences*, *124*, 150–159.

Back, T. (1996). *Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms.* Oxford university press.

Baddeley, M. (2018). *Behavioural economics and finance.* Routledge.

Balakrishnan, V., Khan, S., Fernandez, T., & Arabnia, H. R. (2019). Cyberbullying detection on twitter using big five and dark triad features. *Personality and individual differences*, *141*, 252–257.

Barto, A. G. (1997). Reinforcement learning. In *Neural systems for control* (pp. 7–30). Hoboken, New Jersey, USA.: Elsevier.

Battocchi, K., Dillon, E., Hei, M., Lewis, G., Oka, P., Oprescu, M., & Syrgkanis, V. (2023). Econml: A python package for ml-based heterogeneous treatment effects estimation. *Version 0.14.1*.

Bénabou, R., & Tirole, J. (2016). Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, *30*(3), 141–64.

Ben-David, D., & Sade, O. (2001). Robo-advisor adoption, willingness to pay, and trust—before and at the outbreak of the covid-19 pandemic. *Willingness to Pay, and Trust—Before and at the Outbreak of the COVID-19 Pandemic (May 1, 2001).*

Berger, B., Adam, M., Rühr, A., & Benlian, A. (2021). Watch me improve—algorithm aversion and demonstrating the ability to learn. *Business & Information Systems Engineering*, *63*(1), 55–68.

Berrar, D. (2019). Cross-validation. In S. Ranganathan, M. Gribskov, K. Nakai, & C. Schönbach (Eds.), *Encyclopedia of bioinformatics and computational biology* (p. 542-545). Oxford: Academic Press. doi: https://doi.org/10.1016/B978-0-12-809633-8.20349-X

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, *119*(1), 249–275.

Billari, F. C., Fent, T., Prskawetz, A., & Scheffran, J. (2006). Agent-based computational modelling: an introduction. In *Agent-based computational modelling* (pp. 1–16). Springer.

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4) (No. 4). Springer.

Blueschke, D., Savin, I., & Blueschke-Nikolaeva, V. (2020). An evolutionary approach to passive learning in optimal control problems. *Computational Economics*, *56*(3), 659–673.

Bogert, E., Schecter, A., & Watson, R. T. (2021). Humans rely more on algorithms than social influence as a task becomes more difficult. *Scientific reports*, *11*(1), 1–9.

Borch, K. (1982). Insuring and auditing the auditor. In *Games, economic dynamics, and time series analysis* (pp. 117–126).

Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern recognition*, *30*(7), 1145–1159.

Brandts, J., & Schwieren, C. (2007). Frames and games. In *Developments on experimental economics* (pp. 175–180). Springer.

Brandts, J., & Schwieren, C. (2009). Frames and economic behavior. *Working Paper*.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

Brewer, M. B., & Kramer, R. M. (1986). Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of personality and social psychology*, *50*(3), 543.

Brown, J. N., & Rosenthal, R. W. (1990). Testing the minimax hypothesis: a re-examination of o'neill's game experiment. *Econometrica: Journal of the Econometric Society*, 1065–1081.

Brunner, C., Camerer, C., & Goeree, J. K. (2010). A correction and re-examination of'stationary concepts for experimental 2x2 games'. *American Economic Review, forthcoming*.

Brunner, C., Camerer, C. F., & Goeree, J. K. (2011). Stationary concepts for experimental 2 x 2 games: Comment. *American Economic Review*, *101*(2), 1029–40.

Bruns, B. (2010). Navigating the topology of 2x2 games: an introductory note on payoff families, normalization, and natural order. *arXiv preprint arXiv:1010.4727*.

Bruns, B. (2011). Visualizing the topology of 2x2 games: From prisoner's dilemma to win-win. *In Stony Brook, NY: Game Theory Center*.

Bruns, B. (2015a). *Atlas of 2x2 games: Transforming conflict and cooperation.*

Bruns, B. (2015b). Names for games: locating 2× 2 games. *Games*, *6*(4), 495–520.

Bucklin, R., Lehmann, D., & Little, J. (1998). From decision support to decision automation: A 2020 vision. *Marketing Letters*, *9*, 235–246.

Bullard, J., & Duffy, J. (1998). Learning and the stability of cycles. *Macroeconomic Dynamics*, *2*(1), 22–48.

Burton, J. W., Stein, M.-K., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*, *33*(2), 220–239.

Camerer, C. (2003). Strategizing in the brain. *Science*, *300*(5626), 1673–1675.

Camerer, C. (2011). *Behavioral game theory: Experiments in strategic interaction.* Princeton University Press.

Camerer, C. (2012). The promise of lab-field generalizability in experimental economics: A reply to levitt and list (2007). *Methods of Modern Experimental Economics, Oxford University Press, Forthcoming*.

Camerer, C., & Hua Ho, T. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, *67*(4), 827–874.

Camerer, C., & Weigelt, K. (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica: Journal of the Econometric Society*, 1–36.

Camerer, C. F. (2018). Artificial intelligence and behavioral economics. In *The economics of artificial intelligence: An agenda* (pp. 587–608). University of Chicago Press.

Camerer, C. F., & Loewenstein, G. (2003). Behavioral economics: Past, present, future.

Candrian, C., & Scherer, A. (2022). Rise of the machines: Delegating decisions to autonomous ai. *Computers in Human Behavior*, *134*, 107308.

Castelluccia, C., & Le Métayer, D. (2019). Understanding algorithmic decision-making: Opportunities and challenges.

Castillo, L., & Dorao, C. (2013). Decision-making in the oil and gas projects based on game theory: Conceptual process design. *Energy Conversion and Management*, *66*, 48–55.

Castillo, L., & Dorao, C. A. (2012). Consensual decision-making model based on game theory for lng processes. *Energy conversion and management*, *64*, 387–396.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of economic behavior & organization*, *81*(1), 1–8.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The quarterly journal of economics*, *117*(3), 817–869.

Chen, D. L., Schonger, M., & Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.

Chen, H., Harinen, T., Lee, J.-Y., Yung, M., & Zhao, Z. (2020). Causalml: Python package for causal machine learning. *arXiv preprint arXiv:2002.11631*.

Chen, S.-H., Duffy, J., & Yeh, C.-H. (2005). Equilibrium selection via adaptation: Using genetic programming to model learning in a coordination game. In *Advances in dynamic games* (pp. 571–598). Springer.

Cheng, L., Guo, R., Moraffah, R., Sheth, P., Candan, K. S., & Liu, H. (2022). *Evaluation methods and measures for causal learning algorithms.*

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., & Newey, W. (2017). Double/debiased/neyman machine learning of treatment effects. *American Economic Review*, *107*(5), 261–265.

Cheung, Y. K., & Egan, V. (2021). The hexaco-60, the dark triad and scholastic cheating. *Psychological Reports*, *124*(6), 2774–2794.

Chmura, T., Goerg, S. J., & Selten, R. (2012). Learning in experimental 2×2 games. *Games and Economic Behavior*, *76*(1), 44 - 73.

Chmura, T., Goerg, S. J., & Selten, R. (2014). Generalized impulse balance: an experimental test for a class of 3 x 3 games. *Review of Behavioral Economics*, *1*(1-2), 27–53.

Chmura, T., Kaiser, J., & Pitz, T. (2007). Simulating complex social behaviour with the genetic action tree kernel. *Computational and Mathematical Organization Theory*, *13*(4), 355–377.

Choi, S., Gale, D., & Kariv, S. (2012). Social learning in networks: a quantal response equilibrium analysis of experimental data. *Review of Economic Design*, *16*(2), 135–157.

Chong, L., Zhang, G., Goucher-Lambert, K., Kotovsky, K., & Cagan, J. (2022). Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of ai advice. *Computers in Human Behavior*, *127*, 107018.

Chugunova, M., & Sele, D. (2022). We and it: An interdisciplinary review of the experimental evidence on how humans interact with machines. *Journal of Behavioral and Experimental Economics*, *99*, 101897.

Colman, A. M. (2016). *Game theory and experimental games: The study of strategic interaction.* Elsevier.

Cookson, R. (2000). Framing effects in public goods experiments. *Experimental Economics*, *3*(1), 55–79.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2001). *Introduction to algorithms* (2nd ed.). Cambridge, Massachusetts: The MIT Press.

Costa Jr, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: robust and surprising findings. *Journal of personality and social psychology*, *81*(2), 322.

Crysel, L. C., Crosier, B. S., & Webster, G. D. (2013). The dark triad and risk behavior. *Personality and Individual Differences*, *54*(1), 35-40. doi: https://doi.org/10.1016/j.paid.2012.07.029

Curtis, G. J., Correia, H. M., & Davis, M. C. (2022). Entitlement mediates the relationship between dark triad traits and academic misconduct. *Personality and Individual Differences*, *191*, 111563. doi: https://doi.org/10.1016/j.paid.2022.111563

Curtis, S. R., Basak, A., Carre, J. R., Bošanský, B., Černý, J., Ben-Asher, N., ... Kiekintveld, C. (2021). The dark triad and strategic resource control in a competitive computer game. *Personality and Individual Differences*, *168*, 110343. doi: https://doi.org/10.1016/j.paid.2020.110343

Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. In C. Zhang & Y. Ma (Eds.), *Ensemble machine learning: Methods and applications* (pp. 157–175). Springer New York. doi: 10.1007/978-1-4419-9326-7_5

Davis, J. M., & Heller, S. B. (2017, May). Using causal forests to predict treatment heterogeneity: An application to summer jobs. *American Economic Review*, *107*(5), 546-50. Retrieved from `https://www.aeaweb.org/articles?id=10.1257/aer.p20171000` doi: 10.1257/aer.p20171000

Dawid, H. (1999). *Adaptive learning by genetic algorithms, analytical results and applications to economic models.*

Dawid, H., & Kopel, M. (1998). On economic applications of the genetic algorithm: a model of the cobweb type. *Journal of Evolutionary Economics*, *8*(3), 297–315.

De Bruyn, A., Viswanathan, V., Beh, Y. S., Brock, J. K.-U., & von Wangenheim, F. (2020). Artificial intelligence and marketing: Pitfalls and opportunities. *Journal of*

Interactive Marketing, *51*, 91–105.

Deutchman, P., & Sullivan, J. (2018). The dark triad and framing effects predict selfish behavior in a one-shot prisoner's dilemma. *PloS one*, *13*(9), e0203891.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (Vol. 326). John Wiley & Sons.

Dresher, M. (1962). *A sampling inspection problem in arms control agreements: A game-theoretic analysis* (Tech. Rep.). Rand Corp Santa Monica CA.

Duffy, J. (2006). Agent-based models and human subject experiments. *Handbook of computational economics*, *2*, 949–1011.

Dufwenberg, M., Gächter, S., & Hennig-Schmidt, H. (2011). The framing of games and the psychology of play. *Games and Economic Behavior*, *73*(2), 459–478.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American statistical association*, *56*(293), 52–64.

Eil, D., & Rao, J. M. (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, *3*(2), 114–38.

Ellickson, P. B., Kar, W., & Reeder III, J. C. (2023). Estimating marketing component effects: Double machine learning from targeted digital promotions. *Marketing Science*, *42*(4), 704–728.

Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2021). Artificial intelligence and business value: A literature review. *Information Systems Frontiers*, 1–26.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, 848–881.

Esteves, G. G. L., Oliveira, L. S., de Andrade, J. M., & Menezes, M. P. (2021). Dark triad predicts academic cheating. *Personality and Individual Differences*, *171*, 110513. doi: https://doi.org/10.1016/j.paid.2020.110513

Fawcett, T. (2006). An introduction to roc analysis. *Pattern recognition letters*, *27*(8), 861–874.

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

Fehr, E., Naef, M., & Schmidt, K. M. (2006). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Comment. *American Economic Review*, *96*(5), 1912–1917.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, *114*(3), 817–868.

Fenneman, A., Sickmann, J., Pitz, T., & Sanfey, A. G. (2021). Two distinct and separable processes underlie individual differences in algorithm adherence: Differences in predictions and differences in trust thresholds. *Plos one*, *16*(2), e0247084.

Ferguson, T. S., & Melolidakis, C. (1998). On the inspection game. *Naval Research*

Logistics (NRL), *45*(3), 327–334.

Filiz, I., Judek, J. R., Lorenz, M., & Spiwoks, M. (2021). The tragedy of algorithm aversion.

Filiz, I., Judek, J. R., Lorenz, M., & Spiwoks, M. (2022). Algorithm aversion as an obstacle in the establishment of robo advisors. *Journal of Risk and Financial Management*, *15*(8), 353.

Fink, D., Johnston, A., Strimas-Mackey, M., Auer, T., Hochachka, W. M., Ligocki, S., ... others (2022). A double machine learning trend model for citizen science data. *arXiv preprint arXiv:2210.15524*.

Fisher, R. A. (1936). Design of experiments. *British Medical Journal*, *1*(3923), 554.

Fogel, D. B. (1998). *Artificial intelligence through simulated evolution.* Wiley-IEEE Press.

Fogel, D. B. (2000). What is evolutionary computation? *IEEE spectrum*, *37*(2), 26–32.

Freedman, L. (1998). *The revolution in strategic affairs.* Oxford University Press, USA.

Freedman, L. (2017). *The transformation of strategic affairs* (No. 379). Routledge.

Friedman, D. (1991). Evolutionary games in economics. *Econometrica: Journal of the Econometric Society*, 637–666.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189–1232.

Fudenberg, D., & Tirole, J. (1991). *Game theory.* MIT press.

Gaspars-Wieloch, H. (2014). Modifications of the hurwicz's decision rule. *Central European Journal of Operations Research*, *22*(4), 779–794.

Gaudeul, A., & Giannetti, C. (2023). Trade-offs in the design of financial algorithms. *Available at SSRN 4432707*.

Gilli, M., & Winker, P. (2009). Heuristic optimization methods in econometrics. *Handbook of computational econometrics*, 81–119.

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining explanations: An overview of interpretability of machine learning. In *2018 ieee 5th international conference on data science and advanced analytics (dsaa)* (pp. 80–89).

Gintis, H., et al. (2000). *Game theory evolving: A problem-centered introduction to modeling strategic behavior.* Princeton university press.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627–660.

Glynatsi, N. E., Knight, V., & Lee, T. E. (2018). An evolutionary game theoretic model of rhino horn devaluation. *Ecological Modelling*, *389*, 33–40.

Goette, L., Han, H.-J., & Leung, B. T. K. (2020). Information overload and confirmation bias.

Goldbach, C., Kayar, D., Pitz, T., & Sickmann, J. (2019). Transferring decisions to an algorithm: A simple route choice experiment. *Transportation research part F: traffic psychology and behaviour*, *65*, 402–417.

Goldbach, C., Sickmann, J., Pitz, T., & Zimasa, T. (2022). Towards autonomous public transportation: attitudes and intentions of the local population. *Transportation Research Interdisciplinary Perspectives*, *13*, 100504.

Goldberg, D. E., & Holland, J. H. (1988). Genetic algorithms and machine learning.

Machine learning, *3*(2), 95–99.

Goldberg, L. R. (1990). An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, *59*(6), 1216.

Goldman, A. I. (1971). The individuation of action. *The journal of Philosophy*, *68*(21), 761–774.

Gooding, T. (2014). Modelling society's evolutionary forces. *Journal of Artificial Societies and Social Simulation*, *17*(3), 3.

Goodman, B., & Flaxman, S. (2017). European union regulations on algorithmic decision-making and a "right to explanation". *AI magazine*, *38*(3), 50–57.

Gottlieb, J., & Oudeyer, P.-Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience*, *19*(12), 758–770.

Green, J., White, M. H., et al. (2023). *Machine learning for experiments in the social sciences*. Cambridge University Press.

Greenberg, B. G., Abul-Ela, A.-L. A., Simmons, W. R., & Horvitz, D. G. (1969). The unrelated question randomized response model: Theoretical framework. *Journal of the American Statistical Association*, *64*(326), 520–539.

Guelman, L., Guillén, M., & Pérez-Marín, A. M. (2012). Random forests for uplift modeling: an insurance customer retention case. In *International conference on modeling and simulation in engineering, economics and management* (pp. 123–133).

Guelman, L., Guillén, M., & Pérez-Marín, A. M. (2015). Uplift random forests. *Cybernetics and Systems*, *46*(3-4), 230–248.

Gunning, D. (2017). Explainable artificial intelligence (xai). *Defense advanced research projects agency (DARPA), nd Web*, *2*(2), 1.

Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences*, *18*, 43–49.

Gurvich, V. A., Karzanov, A. V., & Khachivan, L. (1988). Cyclic games and an algorithm to find minimax cycle means in directed graphs. *USSR Computational Mathematics and Mathematical Physics*, *28*(5), 85–91.

Gutierrez, P., & Gérardy, J.-Y. (2017). Causal inference and uplift modelling: A review of the literature. In *International conference on predictive applications and apis* (pp. 1–13).

Hagenbach, J., & Koessler, F. (2022). Selective memory of a psychological agent. *European Economic Review*, 104012.

Hamblin, S. (2013). On the practical usage of genetic algorithms in ecology and evolution. *Methods in Ecology and Evolution*, *4*(2), 184–194.

Hamblin, S., & Hurd, P. L. (2007). Genetic algorithms and non-ess solutions to game theory models. *Animal Behaviour*, *74*(4), 1005–1018.

Harsanyi, J. C., Selten, R., et al. (1988). A general theory of equilibrium selection in games. *MIT Press Books*, *1*.

Haruvy, E., & Ünver, M. U. (2002). Equilibrium selection in repeated b2b matching markets".

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of* *statistical learning: data mining, inference, and prediction* (Vol. 2). Springer.

Haupt, R. L., Haupt, S. E., & Haupt, S. E. (1998). *Practical genetic algorithms* (Vol. 2). Wiley New York.

Hayashi, N., Ostrom, E., Walker, J., & Yamagishi, T. (1999). Reciprocity, trust, and the sense of control: A cross-societal study. *Rationality and society*, *11*(1), 27–46.

Heckathorn, D. D. (1996). The dynamics and dilemmas of collective action. *American sociological review*, 250–277.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, *91*(2), 73–78.

Herm, L.-V., Heinrich, K., Wanner, J., & Janiesch, C. (2022). Stop ordering machine learning algorithms by their explainability! a user-centered investigation of performance and explainability. *International Journal of Information Management*, 102538.

Hines, W. (1987). Evolutionary stable strategies: a review of basic theory. *Theoretical Population Biology*, *31*(2), 195–272.

Hoelzemann, J., & Klein, N. (2021). Bandits in the lab. *Quantitative Economics*, *12*(3), 1021–1051.

Holland, J. H. (1992). Genetic algorithms. *Scientific american*, *267*(1), 66–73.

Holland, J. H., & Miller, J. H. (1991). Artificial adaptive agents in economic theory. *The American economic review*, *81*(2), 365–370.

Holler, M. J. (1993). Fighting pollution when decisions are strategic. *Public Choice*, *76*, 347–356.

Hou, Y. T.-Y., & Jung, M. F. (2021). Who is the expert? reconciling algorithm aversion and algorithm appreciation in ai-supported decision making. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–25.

Hunt, L. T., Rutledge, R. B., Malalasekera, W. N., Kennerley, S. W., & Dolan, R. J. (2016). Approach-induced biases in human information sampling. *PLoS biology*, *14*(11), e2000638.

Hurwicz, L. (1951). The generalized bayes minimax principle: a criterion for decision making under uncertainty. *Cowles Comm. Discuss. Paper Stat*, *335*, 1950.

Isaac, A. G. (2008). Simulating evolutionary games: a python-based introduction. *Journal of Artificial Societies and Social Simulation*, *11*(3), 8.

Jaffray, J.-Y., Jeleva, M., Gains, U., & Paris, C. (2007). Information processing under imprecise risk with the hurwicz criterion. In *5th international symposium on imprecise probability: theories and applications* (pp. 233–242).

Jago, A. S., & Laurin, K. (2022). Assumptions about algorithms' capacity for discrimination. *Personality and Social Psychology Bulletin*, *48*(4), 582–595.

Jauernig, J., Uhl, M., & Walkowitz, G. (2022). People prefer moral discretion to algorithms: algorithm aversion beyond intransparency. *Philosophy & Technology*, *35*(1), 1–25.

Jones, D. N., & Neria, A. L. (2015). The dark triad and dispositional aggression. *Personality and Individual Differences*, *86*, 360-364. doi: https://doi.org/10.1016/j.paid.2015.06.021

Jung, M., & Seiter, M. (2021). Towards a better understanding on mitigating algorithm aversion in forecasting: an experimental study. *Journal of Management Control*, *32*(4), 495–516.

Jussupow, E., Benbasat, I., & Heinzl, A. (2020). Why are we averse towards algorithms? a comprehensive literature review on algorithm aversion.

Kahneman, D. (2011). *Thinking, fast and slow.* United Kingdom: macmillan.

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

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American psychologist*, *39*(4), 341.

Kalai, E., & Lehrer, E. (1993). Rational learning leads to nash equilibrium. *Econometrica: Journal of the Econometric Society*, 1019–1045.

Kaplansky, I. (1945). A contribution to von neumann's theory of games. *Annals of Mathematics*, 474–479.

Kaplansky, I. (1995). A contribution to von neumann's theory of games. ii. *Linear algebra and its applications*, *226*, 371–373.

Kelly, C., Sharot, T., et al. (2021). Individual differences in information-seeking. *Nature communications*, *12*(1), 1–13.

Kirman, A. P., Laisney, F., & Pezanis-Christou, P. (2018). Exploration vs exploitation, impulse balance equilibrium, and a specification test for the el farol bar problem. *Impulse Balance Equilibrium, and a Specification Test for the El Farol Bar Problem (August 16, 2018).*

Knight, V., & Campbell, J. (2018). Nashpy: A python library for the computation of nash equilibria. *Journal of Open Source Software*, *3*(30), 904.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, 33–50.

Kohavi, R., et al. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai* (Vol. 14, pp. 1137–1145).

Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual review of sociology*, *24*(1), 183–214.

Kolokoltsov, V. (2010). Inspection games and crime prevention. *Department of Statistics, University of Warwick*.

Kolokoltsov, V., Passi, H., & Yang, W. (2013). Inspection and crime prevention: an evolutionary perspective. *arXiv preprint arXiv:1306.4219*.

König, P. D., Wurster, S., & Siewert, M. B. (2022). Consumers are willing to pay a price for explainable, but not for green ai. evidence from a choice-based conjoint analysis. *Big Data & Society*, *9*(1), 20539517211069632.

Kool, W., & Botvinick, M. (2018). Mental labour. *Nature human behaviour*, *2*(12), 899–908.

Koza, J. R. (1994). *Genetic programming ii: automatic discovery of reusable programs.* MIT press.

Koza, J. R., et al. (1992). Evolution of subsumption using genetic programming. In *Proceedings of the first european conference on artificial life* (pp. 110–119).

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, *47*(260), 583–621.

Ku, C.-Y. (2020). When ais say yes and i say no: On the tension between ai's decision and human's decision from the epistemological perspectives. *Információs Társadalom*, *19*(4), 61–76.

Kusyk, J., Sahin, C. S., Uyar, M. U., Urrea, E., & Gundry, S. (2011). Self-organization of nodes in mobile ad hoc networks using evolutionary games and genetic algorithms. *Journal of Advanced Research*, *2*(3), 253–264.

Lainidi, O., Karakasidou, E., & Montgomery, A. (2022). Dark triad, impulsiveness and honesty-humility in the prisoner's dilemma game: The moderating role of gender. *Merits*. doi: 10.3390/merits2040027

Lam, S. K., Pitrou, A., & Seibert, S. (2015). Numba: A llvm-based python jit compiler. In *Proceedings of the second workshop on the llvm compiler infrastructure in hpc* (pp. 1–6).

LaValle, S. M., Branicky, M. S., & Lindemann, S. R. (2004). On the relationship between classical grid search and probabilistic roadmaps. *The International Journal of Robotics Research*, *23*(7-8), 673–692.

Lechner, M. (2023). Causal machine learning and its use for public policy. *Swiss Journal of Economics and Statistics*, *159*(1), 1–15.

Lensberg, T., & Schenk-Hoppé, K. R. (2021). Cold play: Learning across bimatrix games. *Journal of Economic Behavior & Organization*, *185*, 419–441.

Leszek, R. (2008). *Computational intelligence: Methods and techniques.* Springer Press.

Leung, B. T. K. (2020). Limited cognitive ability and selective information processing. *Games and Economic Behavior*, *120*, 345–369.

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, *76*(2), 149-188. doi: https://doi.org/10.1006/obhd.1998.2804

Liashchynskyi, P., & Liashchynskyi, P. (2019). Grid search, random search, genetic algorithm: a big comparison for nas. *arXiv preprint arXiv:1912.06059*.

Liaw, A., Wiener, M., et al. (2002). Classification and regression by randomforest. *R news*, *2*(3), 18–22.

Liu, Y., Tang, W., He, J., Liu, Y., Ai, T., & Liu, D. (2015). A land-use spatial optimization model based on genetic optimization and game theory. *Computers, Environment and Urban Systems*, *49*, 1–14.

Logg, J. M. (2017). Theory of machine: When do people rely on algorithms? *Harvard Business School working paper series# 17-086*.

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103.

Lopez, R., Calvo, J. L., & Torre, I. D. L. (2022). Analyzing the relation between anchoring and dark triad personality traits in the prisoner's and traveler's dilemmas. *Studies of Applied Economics*. doi: 10.25115/eea.v40i2.5760

Lucasius, C. B., & Kateman, G. (1993). Understanding and using genetic algorithms part 1. concepts, properties and context. *Chemometrics and intelligent laboratory systems*, *19*(1), 1–33.

Lundberg, I., Brand, J. E., & Jeon, N. (2022). Researcher reasoning meets computational capacity: Machine learning for social science. *Social science research*, *108*, 102807.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., ... Lee, S.-I. (2020). From local explanations to global understanding with explainable ai for trees. *Nature machine intelligence*, *2*(1), 56–67.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.

Macedo, L. L., Godinho, P., & Alves, M. J. (2020). A comparative study of technical trading strategies using a genetic algorithm. *Computational Economics*, *55*(1), 349–381.

Machlanski, D., Samothrakis, S., & Clarke, P. (2023). Hyperparameter tuning and model evaluation in causal effect estimation. *arXiv preprint arXiv:2303.01412*.

Mahmud, H., Islam, A. N., Ahmed, S. I., & Smolander, K. (2022). What influences algorithmic decision-making? a systematic literature review on algorithm aversion. *Technological Forecasting and Social Change*, *175*, 121390.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50–60.

Manson, S. M. (2006). Bounded rationality in agent-based models: experiments with evolutionary programs. *International Journal of Geographical Information Science*, *20*(9), 991–1012.

Maschler, M. (1966). A price leadership method for solving the inspector's non-constant-sum game. *Naval research logistics quarterly*, *13*(1), 11–33.

Maschler, M., Solan, E., & Zamir, S. (2013). *Game theory*. Cambridge University Press. doi: 10.1017/CBO9780511794216

Mauersberger, F., & Nagel, R. (2018). Levels of reasoning in keynesian beauty contests: a generative framework. In *Handbook of computational economics* (Vol. 4, pp. 541–634). Elsevier.

Mauersberger, F., Nagel, R., & Bühren, C. (2020). Bounded rationality in keynesian beauty contests: a lesson for central bankers? *Economics*, *14*(1), 20200016.

McCarthy, J. (2007). What is artificial intelligence.

McKelvey, R. D., McLennan, A. M., & Turocy, T. L. (2023). Gambit: Software tools for game theory.

McKelvey, R. D., & Palfrey, T. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, *10*(1), 6-38. Retrieved from `https://EconPapers.repec.org/RePEc:eee:gamebe:v:10:y:1995:i:1:p:6-38`

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence*, *267*, 1–38.

Mitchell, M. (1998). *An introduction to genetic algorithms*. MIT press.

Mizuguchi, T., & Sawamura, S. (2023). Machine learning-based causal models for predicting the response of individual patients to dexamethasone treatment as prophylactic antiemetic. *Scientific Reports*, *13*(1), 1–10.

Moffatt, P. (2020). *Experimetrics: Econometrics for experimental economics*. Bloomsbury Publishing.

Mohamed, F. A., & Koivo, H. N. (2011). Multiobjective optimization using modified game theory for online management of microgrid. *European Transactions on Electrical Power*, *21*(1), 839–854.

Moshagen, M., Hilbig, B. E., & Zettler, I. (2018). The dark core of personality. *Psychological review*, *125*(5), 656.

Moshagen, M., Zettler, I., & Hilbig, B. E. (2020). Measuring the dark core of personality. *Psychological Assessment*, *32*(2), 182.

Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87–106.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Nash Jr, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, *36*(1), 48–49.

Nosenzo, D., Offerman, T., Sefton, M., & Van der Veen, A. (2010). *Inducing good behavior: Bonuses versus fines in inspection games* (Tech. Rep.). CeDEx discussion paper series.

Nosenzo, D., Offerman, T., Sefton, M., & van der Veen, A. (2014). Encouraging compliance: Bonuses versus fines in inspection games. *The Journal of Law, Economics, & Organization*, *30*(3), 623–648.

Nosenzo, D., Offerman, T., Sefton, M., & van der Veen, A. (2016). Discretionary sanctions and rewards in the repeated inspection game. *Management Science*, *62*(2), 502–517.

Ockenfels, A., & Selten, R. (2005). Impulse balance equilibrium and feedback in first price auctions. *Games and Economic Behavior*, *51*(1), 155–170.

Osborne, M. J., & Rubinstein, A. (1998). Games with procedurally rational players. *American Economic Review*, 834–847.

Osborne, M. J., & Rubinstein, A. (2003). Sampling equilibrium, with an application to strategic voting. *Games and Economic Behavior*, *45*(2), 434–441.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*(6), 556-563. doi: https://doi.org/10.1016/S0092-6566(02)00505-6

Pažek, K., & Rozman, Č. (2009). Decision making under conditions of uncertainty in agriculture: a case study of oil crops. *Poljoprivreda*, *15*(1), 45–50.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pereira, R. L., Souza, D. L., Mollinetti, M. A. F., Neto, M. T. S., Yasojima, E. K. K., Teixeira, O. N., & De Oliveira, R. C. L. (2020). Game theory and social interaction for selection and crossover pressure control in genetic algorithms: An empirical analysis to real-valued constrained optimization. *IEEE Access*, *8*, 144839–144865.

Périaux, J., Chen, H., Mantel, B., Sefrioui, M., & Sui, H. (2001). Combining game theory and genetic algorithms with application to ddm-nozzle optimization problems. *Finite elements in analysis and design*, *37*(5), 417–429.

Petitet, P., Attaallah, B., Manohar, S. G., & Husain, M. (2021). The computational cost of active information sampling before decision-making under uncertainty. *Nature Human Behaviour*, *5*(7), 935–946.

Pitz, T., Chmura, T., et al. (2005). *Genetic action trees a new concept for social and economic simulation* (Tech. Rep.). Germany: University Library of Munich.

Plonsky, O., Apel, R., Ert, E., Tennenholtz, M., Bourgin, D., Peterson, J. C., ... others (2019). Predicting human decisions with behavioral theories and machine learning. *arXiv preprint arXiv:1904.06866*.

Plous, S. (1993). The nuclear arms race: prisoner's dilemma or perceptual dilemma? *Journal of Peace Research*, *30*(2), 163–179.

Powers, D. M. (2020). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.

Prahl, A., & Van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting*, *36*(6), 691–702.

Price, T. C. (1997). Using co-evolutionary programming to simulate strategic behaviour in markets. *Journal of evolutionary economics*, *7*(3), 219–254.

Puerto, J., Mármol, A., Monroy, L., & Fernández, F. (2000). Decision criteria with partial information. *International Transactions in Operational Research*, *7*(1), 51–65.

Radcliffe, N., & Surry, P. (1999). Differential response analysis: Modeling true responses by isolating the effect of a single action. *Credit Scoring and Credit Control IV*.

Radcliffe, N. J., & Surry, P. D. (2011). Real-world uplift modelling with significance-based uplift trees. *White Paper TR-2011-1, Stochastic Solutions*, 1–33.

Rauhut, H. (2009). Higher punishment, less control? experimental evidence on the inspection game. *Rationality and Society*, *21*(3), 359–392.

Rechenberg, I. (1978). Evolutionsstrategien. In *Simulationsmethoden in der medizin und biologie* (pp. 83–114). Springer.

Reeves, C., & Rowe, J. E. (2002). *Genetic algorithms: principles and perspectives: a guide to ga theory* (Vol. 20). Springer Science & Business Media.

Reschke, C. H., et al. (2001). Evolutionary perspectives on simulations of social systems. *Journal of Artificial Societies and Social Simulation*, *4*(4), 10.

Rey-Biel, P. (2008). Inequity aversion and team incentives. *scandinavian Journal of Economics*, *110*(2), 297–320.

Riechmann, T. (1998). *Genetic algorithms and economic evolution* (Tech. Rep.). Diskussionsbeitrag.

Riechmann, T. (2001). Genetic algorithm learning and evolutionary games. *Journal of Economic Dynamics and Control*, *25*(6-7), 1019–1037.

Riechmann, T. (2002). Genetic algorithm learning and economic evolution. In *Evolutionary computation in economics and finance* (pp. 45–60). Springer.

Robbins, H. (1952). Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, *58*(5), 527–535.

Robbins, H., & Monro, S. (1951). A stochastic approximation method. *The annals of mathematical statistics*, 400–407.

Robinson, D., & Goforth, D. (2005). *The topology of the 2x2 games: a new periodic table*

(Vol. 3). Psychology Press.

Roca, C. P., Cuesta, J. A., & Sanchez, A. (2009). Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. *Physics of life reviews*, *6*(4), 208–249.

Rogoza, R., Kowalski, C. M., Saklofske, D. H., & Schermer, J. A. (2022). Systematizing dark personality traits within broader models of personality. *Personality and Individual Differences*, *186*, 111343. doi: https://doi.org/10.1016/j.paid.2021.111343

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied*, *80*(1), 1.

Rowe, J. E. (2015). Genetic algorithms. In J. Kacprzyk & W. Pedrycz (Eds.), *Springer handbook of computational intelligence* (pp. 825–844). Berlin, Heidelberg: Springer.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, *66*(5), 688.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, *323*(6088), 533–536.

Russell, B. (1959). *Common sense and nuclear warfare.*

Russell, S. J. (2010). *Artificial intelligence a modern approach* (4th ed.). London: Pearson Education, Inc.

Rzepakowski, P., & Jaroszewicz, S. (2012). Decision trees for uplift modeling with single and multiple treatments. *Knowledge and Information Systems*, *32*, 303–327.

Saltık, Ö., Söyü, R., Değirmen, S., Şengönül, A., et al. (2023). Predicting loss aversion behavior with machine-learning methods. *Humanities and Social Sciences Communications*, *10*(1), 1–14.

Samuelson, L. (2005). Economic theory and experimental economics. *Journal of Economic Literature*, *43*(1), 65–107.

Santos, A. C. (2011). Behavioural and experimental economics: are they really transforming economics? *Cambridge journal of economics*, *35*(4), 705–728.

Savin, I., & Blueschke, D. (2016). Lost in translation: Explicitly solving nonlinear stochastic optimal control problems using the median objective value. *Computational Economics*, *48*(2), 317–338.

Savin, I., Blueschke, D., & Blueschke-Nikolaeva, V. (2018). Slow and steady wins the race: Approximating nash equilibria in nonlinear quadratic tracking games steter tropfen höhlt den stein: Approximation von nash gleichgewichten in nicht-linearen dynamischen spielen. *Jahrbücher für Nationalökonomie und Statistik*, *238*(6), 541–569.

Savin, I., & Egbetokun, A. (2016). Emergence of innovation networks from r&d cooperation with endogenous absorptive capacity. *Journal of Economic Dynamics and Control*, *64*, 82–103.

Schmertmann, C. P. (1996). Functional search in economics using genetic programming. *Computational Economics*, *9*(4), 275–298.

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? sex differences in big five personality traits across 55 cultures. *Journal of*

personality and social psychology, *94*(1), 168.
Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In *9th python in science conference.*
Sefrioui, M., & Perlaux, J. (2000). Nash genetic algorithms: examples and applications. In *Proceedings of the 2000 congress on evolutionary computation. cec00 (cat. no. 00th8512)* (Vol. 1, pp. 509–516).
Selten, R., & Buchta, J. (1994). *Experimental sealed bid first price auctions with directly observed bid functions.* Univ., Sonderforschungsbereich 303.
Selten, R., & Chmura, T. (2008). Stationary concepts for experimental 2x2-games. *American Economic Review*, *98*(3), 938–66.
Selten, R., & Wooders, M. H. (2001). Cyclic games: an introduction and some examples. *Games and Economic Behavior*, *34*(1), 138–152.
Shapley, L. S., et al. (1953). A value for n-person games.
Sharan, N. N., & Romano, D. M. (2020). The effects of personality and locus of control on trust in humans versus artificial intelligence. *Heliyon*, *6*(8), e04572.
Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *Journal of economic perspectives*, *17*(1), 83–104.
Simpson, B. (2004). Social values, subjective transformations, and cooperation in social dilemmas. *Social Psychology Quarterly*, *67*(4), 385–395.
Skyrms, B. (2004). *The stag hunt and the evolution of social structure.* Cambridge University Press.
Slowik, A., & Kwasnicka, H. (2020). Evolutionary algorithms and their applications to engineering problems. *Neural Computing and Applications*, *32*(16), 12363–12379.
Smith, K., Emerson, D., Haight, T., & Wood, B. (2022). An examination of online cheating among business students through the lens of the dark triad and fraud diamond. *Ethics & Behavior*. doi: 10.1080/10508422.2022.2104281
Smith, V. L. (1991). *Papers in experimental economics.* Cambridge University Press.
Snijders, C., Conijn, R., de Fouw, E., & van Berlo, K. (2023). Humans and algorithms detecting fake news: Effects of individual and contextual confidence on trust in algorithmic advice. *International Journal of Human–Computer Interaction*, *39*(7), 1483–1494.
Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information processing & management*, *45*(4), 427–437.
Sołtys, M., Jaroszewicz, S., & Rzepakowski, P. (2015). Ensemble methods for uplift modeling. *Data mining and knowledge discovery*, *29*, 1531–1559.
Spears, W. M., Jong, K. A. D., Bäck, T., Fogel, D. B., & Garis, H. d. (1993). An overview of evolutionary computation. In *European conference on machine learning* (pp. 442–459).
Stachl, C., Hilbert, S., Au, J.-Q., Buschek, D., De Luca, A., Bischl, B., ... Bühner, M. (2017). Personality traits predict smartphone usage. *European Journal of Personality*, *31*(6), 701–722.
Storn, R., & Price, K. (1997). Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, *11*(4),

341–359.

Sundararajan, M., & Najmi, A. (2020, 13–18 Jul). The many shapley values for model explanation. In H. D. III & A. Singh (Eds.), *Proceedings of the 37th international conference on machine learning* (Vol. 119, pp. 9269–9278). PMLR. Retrieved from `https://proceedings.mlr.press/v119/sundararajan20b.html`

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction.* Cambridge Massachussets: MIT press.

Tanimoto, J. (2015). *Fundamentals of evolutionary game theory and its applications.* Springer.

Tavoni, A. (2009). Incorporating fairness motives into the impulse balance equilibrium and quantal response equilibrium concepts: an application to 2x2 games.

Thaler, R. H. (2016). Behavioral economics: Past, present, and future. *American Economic Review*, *106*(7), 1577–1600.

Timmermans, E., Caluwé, E. D., & Alexopoulos, C. (2018). Why are you cheating on tinder? exploring users' motives and (dark) personality traits. *Computers in Human Behavior*. doi: 10.1016/j.chb.2018.07.040

Trocin, C., Mikalef, P., Papamitsiou, Z., & Conboy, K. (2021). Responsible ai for digital health: a synthesis and a research agenda. *Information Systems Frontiers*, 1–19.

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

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *science*, *211*(4481), 453–458.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. doi: 10.1038/s41592-019-0686-2

Vlačić, B., Corbo, L., e Silva, S. C., & Dabić, M. (2021). The evolving role of artificial intelligence in marketing: A review and research agenda. *Journal of Business Research*, *128*, 187–203.

Von Neumann, J., & Morgenstern, O. (1953). *Theory of games and economic behavior.* Princeton university press.

von Stengel, B. (2007). Equilibrium computation for two-player games in strategic and extensive form. *Algorithmic game theory*, 53–78.

von Wedel, P., & Hagist, C. (2022). Physicians' preferences and willingness to pay for artificial intelligence-based assistance tools: a discrete choice experiment among german radiologists. *BMC Health Services Research*, *22*(1), 1–14.

Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, *113*(523), 1228–1242.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, *60*(309), 63–69.

Wasserbacher, H., & Spindler, M. (2022). Machine learning for financial forecasting, planning and analysis: recent developments and pitfalls. *Digital Finance*, *4*(1), 63–

88.

Weibull, J. W. (1997). *Evolutionary game theory.* MIT press.

Weisberg, Y. J., DeYoung, C. G., & Hirsh, J. B. (2011). Gender differences in personality across the ten aspects of the big five. *Frontiers in psychology*, 178.

Widger, J., & Grosu, D. (2008). Computing equilibria in bimatrix games by parallel support enumeration. In *2008 international symposium on parallel and distributed computing* (pp. 250–256).

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the united states and japan. *Motivation and emotion*, *18*(2), 129–166.

Yang, G. (2017). Game theory-inspired evolutionary algorithm for global optimization. *Algorithms*, *10*(4), 111.

Zhang, J., Paulhus, D. L., & Ziegler, M. (2019). Personality predictors of scholastic cheating in a chinese sample. *Educational Psychology*. doi: 10.1080/01443410.2018.1502414

Zhang, Y., Chen, X., et al. (2020). Explainable recommendation: A survey and new perspectives. *Foundations and Trends® in Information Retrieval*, *14*(1), 1–101.