The Behavioral Economics of Artificial Intelligence: Lessons from Experiments with Computer Players

Christoph March

Working Paper No. 154

November 2019

BERG Working Paper Series

Bamberg Economic Research Group
Bamberg University
Feldkirchenstraße 21
D-96052 Bamberg
Telefax: (0951) 863 5547
Telephone: (0951) 863 2687
felix.stuebben@uni-bamberg.de
http://www.uni-bamberg.de/vwl/forschung/berg/

ISBN 978-3-943153-75-0

Redaktion:
Dr. Felix Stübben*

* felix.stuebben@uni-bamberg.de

The Behavioral Economics of Artificial Intelligence: Lessons from Experiments with Computer Players

Christoph March[‡]

This version: October 21, 2019

Abstract

Artificial intelligence (AI) is starting to pervade the economic and social life rendering strategic interactions with artificial agents more and more common. At the same time, experimental economic research has increasingly employed computer players to advance our understanding of strategic interaction in general. What can this strand of research teach us about an AI-shaped future? I review 90 experimental studies using computer players. I find that, in a nutshell, humans act more selfishly and more rational in the presence of computer players, and they are often able to exploit these players. Still, many open questions prevail.

Keywords: Experiment; Robots; Computer players; Survey

JEL classification: C90; C92; O33

[‡]Technische Universität München, Arcisstraße 21, D–80333 Munich, Germany; University of Bamberg, Department of Economics, Feldkirchenstraße 21, D–96047 Bamberg, Germany; and CESifo, Poschingerstraße 5, D–81679 Munich, Germany. Email: christoph.march@tum.de.

1 Introduction

Artificial intelligence (AI) is starting to pervade the social life. Driven by a drastic surge in computer power and the availability of massive and increasing amounts of data, *autonomous, learning systems* are advancing into more and more domains.[1] These include, inter alia, financial transactions (Hendershott et al., 2011, Brogaard et al., 2014), industrial production (Acemoglu and Restrepo, 2019), fraud detection (Raj and Annie, 2011), recruitment (Upadhyay and Khandelwal, 2018), medicine (Topol, 2019), and autonomous driving.

The diffusion of artificial intelligence is set to fundamentally transform our economies and societies. Agrawal et al. (2019) collects a series of papers outlining the economic changes it may bring along, and the novel questions this raises for economists in various fields. In particular, Camerer (2019) points out three implications for behavioral economics: First, artificial intelligence may be used to identify behavioral variables that affect behavior. Second, difficulties in implementing artificial intelligence may help us understand common limitations of human cognition. And finally, behavioral economics is necessary to understand and predict how artificial intelligence may overcome and exploit human limitations.

Camerer's final point deserves further emphasis. In strategic interactions between humans, many deviations from the predictions of equilibrium models stem from the agents' inability to assess others' behavior. A common finding is that humans tend to underestimate the rationality of others which leads them to e.g. underappreciate the effects of changes in the strategic environment, or to learn too little from others' actions (see e.g Eyster, 2019). As artificial intelligence becomes pervasive in various fields, strategic interaction between humans and artificial agents will become more and more common. It is yet unclear how the results on (the limits of) strategic sophistication in human-human-interactions translate to human-computer-interactions. Will humans ascribe too much or too little rationality to artificially intelligent agents? Answers to this and related questions are at the core of understanding and predicting (changes in) economic behavior in the presence of artificial intelligence.[2]

In this paper, I posit that experiments employing computer players have a lot to

[1] In this paper, I use the term "artificial intelligence" to refer to technical systems that are able – to some degree – to collect, combine, and process information, derive conclusions from it, and act upon these conclusions. Currently, many of these systems rely on a form of machine learning, and are highly domain-specific i.e. only perform well under very restricted circumstances. I refrain from discussions on the distinction between such "weak AI" and a "strong AI" that is sufficiently flexible to perform well across a multitude of domains, and therefore exhibits a form of intelligence commonly associated with humans only (see Searle, 1980).

[2] The coexistence and strategic interaction of humans and artificial agents is unlikely to be a transient stage. First, digital technologies diffuse slowly through the economy and society (Andrews and Timiliotis, 2018). Second, labor market policies are currently designed with the aim of enabling humans to work alongside artificially intelligent agents (see e.g. OECD, 2019b).

contribute to answering the above questions. By comparing subjects' behavior in strategic interactions with computer players and with other human subjects, such experiments may identify behavioral regularities that carry over from human-human-interaction to human-computer-interaction, and they may uncover new regularities which potentially emerge. Moreover, I show that experimental economists have employed computer players for over five decades and I review 90 such studies. Though computer players in experimental studies have largely served as a methodological tool to further increase experimental control, the existing experiments already deliver valuable insights into the changes in strategic behavior when (some) subjects are replaced by computers. In addition, more recent studies have deliberately focused on human-computer-interaction. I provide an overview of the literature in terms of topics, aims, and methodological choices, and I derive conclusions from this literature.

Four stylized facts emerge: First, behavior often changes when human opponents are replaced by computer players. Second, subjects generically behave more selfish and more rational when interacting with computers. Third, subjects often learn to exploit computer players, even if the latter do not follow a fixed strategy but are responsive to the subjects' choices, and if subjects possess little information about computers a priori. Fourth, there are limits to exploitation, and sophisticated algorithms are able to outperform human subjects in certain environments.

Conversely, my review also shows that many questions have been left unanswered by this strand of the experimental literature. This is mostly due to the fact that few of the studies have been explicitly designed to study changes in strategic behavior when (some) subjects are replaced by computer players. I elaborate on some of those open issues. I contend that experimentalists have a lot to contribute to the economics of artificial intelligence, and to ongoing debates regarding e.g. the transparency and explainability of AI (OECD, 2019a, European Commission, 2019).

The paper proceeds as follows: In the next section, I describe how I have constructed the database of studies. In Section 3, I give an overview of the studies by topic. I discuss the different aims of using computer players and further methodical aspects in Sections 4 and 5, respectively. Finally, I conclude in Section 6 with a discussion of open questions for future research. The appendix contains an overview and classification of all the experimental studies using computer players I was able to find.

2 Searching for Experiments with Computer Players

I have searched for experiments employing computer players (CP) in three different ways: First, I conducted a search in the *Web of Science* database for papers containing both the word "experiment" and a term from the set

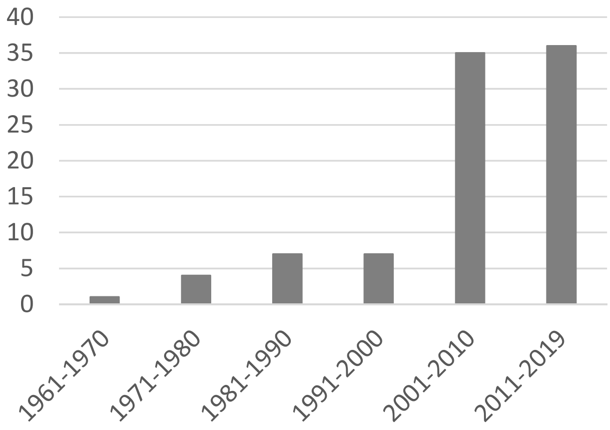
$$\{\text{computer}, \text{computerized}, \text{virtual}, \text{artificial}, \text{programmed}, \text{simulated}\} \times \{\text{players}, \text{agents}, \text{opponent}, \text{partner}\} \cup \{\text{robot}, \text{machine}\}$$

in the set of words used in the title, abstract, or keywords. Second, I have searched among the references given in these papers. Finally, I have done a backwards search for additional studies citing those papers.

One result from this search is that experiments with human and CP have been employed in various fields other than economics, including most notably computer science, neuroscience, and psychology. The literature is therefore large. My focus in this paper is on the strategic interaction between humans and computers, and the implications for the field of economics. Furthermore, I focus on experiments that adhere to the standards of experimental economics. I therefore make the following restrictions: First, I exclude experiments on human-machine-interaction in which strategic interdependencies are either absent or not addressed (e.g. interaction with and optimal design of chatbots or avatars), or which are not grounded in economic theory (e.g. computer games, online shopping). Second, I exclude experiments which do not explicitly refer to CP, but have participants interact with the experimenter (e.g. Neral and Ochs, 1992), inform participants that they are matched with subjects from a previous session (e.g. Shapiro, 2009, Cox et al., 2017), or transform (part of) the game into an individual decision problem using chance moves (e.g. Anderhub et al., 2002, Charness and Levin, 2009). Finally, I exclude papers that use deception. Indeed, several papers do not tell subjects about their interaction with CP, and go to some lengths to avoid making subjects suspicious (e.g. Roth and Schoumaker, 1983, Harrison and McCabe, 1996, Shachat and Swarthout, 2012).

Overall, I have collected 90 studies which satisfy the above criteria. A complete overview of the studies is provided in Appendix 1. Figure 1 depicts the distribution across decades. Though the use of CP is at least as old as the experimental method, there is a clear surge since the turn of the century. In the next section, I review the studies by topic. In sections 4 and 5, I discuss, respectively, the why and how of CP in experiments.

FIGURE 1
Studies Employing Computer Players Across Decades



3 Topics

3.1 Mixed-Strategy Nash Equilibrium

Some of the earliest experimental studies employing CP use them to investigate behavior in two-player zero-sum (or strictly competitive) games. Early examples are Messick (1967) and Fox (1972). More recently, this line of research has been taken up by neuroscientists (e.g. West and Lebiere, 2001, Spiliopoulos, 2013). As those games usually do not possess a Nash equilibrium in pure strategies, the equilibrium prediction entails the play of mixed strategies chosen such that each player is exactly indifferent between all strategies in the mix. The questions whether and how mixed strategy Nash equilibria (MSNE) emerge has spurred a large theoretical and experimental literature in economics and psychology (see e.g Camerer, 2003, Chapter 3).

The main issue here is that the equilibrium requires mixed strategies and beliefs to be exactly coordinated in the sense that even the tiniest deviation by either player (actual or believed) makes the equilibrium strategy for the other player strictly suboptimal. At the same time, a player has no incentive to adhere exactly to her equilibrium strategy even if the other player does, as each strategy in the mixture yields exactly the same expected payoff. In addition, learning models which converge to a MSNE (like fictitious play) often do so only at the population level (Fudenberg and Levine, 1998). The above properties raise several problems for the experimental analysis of MSNE. In particular, unless all other subjects adhere to the equilibrium strategies, one can conclude very little about an individual subject's ability to play MSNE. And though some experimental studies find a convergence to MSNE at the population level, this is a long way from individuals learning to play MSNE.

The use of CP enables the experimenter to gain control about some of these issues by fixing equilibrium play for one party, and investigating the responses of subjects in the other role. Moreover, they enable the experimenter to analyze whether subjects recognize and exploit deviations from equilibrium by programming CP to play suboptimal strategies, and how subjects respond to naive and sophisticated learning dynamics that CP adhere to. Both approaches provide important tests of the dynamic view of MSNE as the outcome of learning. In most studies investigating play in zero-sum games using CP, subjects are informed very little about the algorithm of CP, but they are given plenty learning opportunities as many repetitions are conducted (100 or more).

Three findings emerge in this literature:[3] First, subjects learn to exploit suboptimal play by CP. This suggests that subjects grope for optimality, and it supports the dynamic view of MSNE as the only steady state of such a process. Second, subjects also learn to exploit learning algorithms even absent comprehensive prior knowledge, and even if these algorithms are rather sophisticated (see e.g. Coricelli, 2005). This places some doubt on the empirical significance of convergence results for simple learning models. Finally, the evidence on equilibrium play when CP are programmed to implement the equilibrium strategy is rather mixed, which seems unsurprising given that there is no clear advantage from adhering to the MSNE strategy if CP need not be rendered indifferent.

3.2 Prisoner's Dilemma

A second strand of the early research studies behavior in the prisoner's dilemma (PD). Indeed, the studies collected in Oskamp (1971) suggest that this topic has been investigated even earlier using CP. This early literature shows that cooperation of subjects increases gradually with the propensity of CP to cooperate, and that CP playing contingent strategies (like-tit-for-tat) are much better at evoking concurrent cooperation by subjects than CP with a fixed propensity to cooperate, or play between subjects.

Later studies have largely used CP to investigate the impact of beliefs on cooperation in the repeated PD. In particular, Andreoni and Miller (1993) test Kreps et al.'s (1982) theory of rational cooperation in the presence of type uncertainty. In their experimental setting, each subject is, with a given probability, matched with a CP playing tit-for-tat. Their findings confirm the sequential equilibrium reputation hypothesis of Kreps et al. (1982): Increasing the likelihood of meeting a CP increases subjects beliefs, and thus their propensity to cooperate. Furthermore, cooperation in sessions with a 50% chance of meeting a CP is larger than in sessions without CP with stranger or partner matching of subjects.

Relatedly, Duffy and Xie (2016) test a theory of contagious cooperation in a population of anonymous players, repeatedly matched at random to play an n-player PD. The theory predicts, counter-intuitively, that cooperation is easier to sustain in larger populations. Duffy and Xie confirm this result for populations consisting of CP and only one subject. In

[3]See the appendix for the complete list of papers.

further treatments, they lower the number of CP and find that this considerably weakens the support for the theory. Accordingly, strategic uncertainty tends to hinder cooperation.

3.3 Public Good Games

Despite its relation with the PD, computer players have largely been used for a different purpose in the public good game (PGG).[4] In a nutshell and leaving aside the different variants, subjects in a laboratory PGG are endowed with a budget which they can keep or contribute to a public good. Similar to the PD, individual and collective preferences fall apart in the PGG: Whereas contributing zero is a dominant strategy for each subject, the efficient outcome requires all (or some) subjects to contribute fully to the public good. A major focus of the experimental literature on PGG has been the question why cooperation emerges in such environments. The two main explanations posit that subjects are either "confused", or contribute due to social motives such as altruism or reciprocity (see e.g. Andreoni, 1995).

A literature starting with Houser and Kurzban (2002) has attempted to disentangle the two explanations using CP. Typically, subjects in these experiments interact with CP whose choices are predetermined by the choices of subjects in a previous experiment. As a consequence, multi-round social motives like reciprocity cannot affect choices since a subject's contributions have no impact on the contributions of CP. Moreover, the impact of one-round social motives like altruism is severely diminished, if earnings of CP are not paid to other subjects.[5] The recurrent finding is that contributions are around 50% lower in sessions with CP which suggests that confusion accounts for around one-half of the contributions observed in standard PGG experiments without CP.[6]

One take-away from this literature is that CP reduce the impact of social motives, inducing subjects to act more selfishly.

3.4 Market Experiments

The emergence of competitive market behavior is among the oldest and most studied topics of experimental economics. In a typical market experiment, sellers are assigned units of a good with known values to them (or "production costs"), and buyers are assigned valuations for one or several units of the good and experimental currency to buy these units. In his seminal paper, Smith (1962) demonstrates that the competitive price quickly arises in experimental markets organized as a *double auction* in which both buyers

[4] An exception is Guillen et al. (2010) who investigate how the threat of meeting a CP playing a grim trigger strategy in a PGG increases cooperation.

[5] This leaves open the possibility of altruism towards the experimenter (who gets to keep the earnings of CP). However, the experimental literature has largely ignored this possibility.

[6] Ferraro and Vossler (2010) and Yamakawa and Okano (2016) show that the impact of confusion can be further reduced by simplifying the game and the instructions.

and sellers may repeatedly post and accept offers (bids and asks, respectively). Moreover, this holds even under minimal information about the market conditions. Subsequently, a large literature has examined the robustness of this convergence result with respect to e.g. the market institution, demand and supply conditions, and extensions of the strategic environment such as costly information acquisition, advertising, or entry and exit opportunities.

In particular, a series of papers examines behavior under the *posted-offer market institution*. In this two stage procedure, sellers first post prices and quantities while buyers are placed in a waiting loop. After all sellers have posted their offers, buyers are approached in random order and given the chance to make purchases at the posted prices. In contrast to double auction markets, buyers are rather passive under the posted-offer institution, and public information is confined to a single price quote for each seller which may hamper convergence to the competitive equilibrium. Since the focus in posted-offer market experiments is on seller behavior, and irrational buyer behavior may act as a serious confound, buyers were frequently replaced by CP who are fully demand-revealing, i.e. buy at any price below their valuation.

Posted-offer markets with CP as buyers have been employed e.g. to study market contestability (Coursey et al., 1984, Harrison and McKee, 1985, Brown Kruse, 1991), predatory pricing (Isaac and Smith, 1985), asymmetric market power (Davis and Williams, 1986, Cason and Williams, 1990, Davis and Williams, 1991), and the effect of nonstationarities (Davis et al., 1993). Similar considerations motivate various studies which examine markets with information search and advertising using CP as buyers (Cason and Friedman, 2002, Cason and Datta, 2006, Morgan et al., 2006) as well as the two studies by Kalayci and Potter (2011) and Kalayci (2015) who use CP to isolate the use and impact of price complexities.[7] The general conclusion from these studies is that even infrequent deviations from rational behavior by human buyers – assumed away through the use of CP – can have a substantial impact on the market outcome as they affect sellers' expectations and behavior. In standard posted-offer markets, buyers sometimes withhold demand at high prices which has a disciplining effect on sellers, but weakens e.g. market contestability (see also Brown Kruse, 2008). In search-and-advertising studies, aggressive advertising by sellers has little impact on prices with CP, but causes large deviations from equilibrium with human buyers. Overall, these studies show that CP may evoke distinctly different market outcomes.

One strand of the literature deserves special emphasis: For a long time, both economists (e.g. Gode and Sunder, 1993) and computer scientists (e.g. Tesauro and Das, 2001) have designed trading algorithms for CP, and examined how markets populated solely by these CP evolve. Recent experiments explicitly investigate the interaction between such CP

[7] Kalayci (2015) also implements boundedly rational CP.

and human subjects. In a continuous double auction environment, Das et al. (2001) and De Luca and Cliff (2011) find that sophisticated CP are able to outperform subjects. Conversely, Duersch et al. (2010) show that subjects are able to exploit and outperform many simple learning models like fictitious play in a standard oligopoly experiment where the buyer side is collapsed into a demand function. However, humans are unable to beat a simple imitation algorithm. This strand of the literature is an interesting and important avenue for future research.

3.5 Experimental Asset Markets

In an experimental asset market, subjects repeatedly trade units of an asset in a market institution (usually the double auction) where each unit pays a (random) dividend at the end of each trading period. As subjects know the number of periods and the (average) dividend per period, the fundamental value of the asset is known to subjects in each period. Hence, no trade should occur under common knowledge of rationality and risk-neutrality. In contrast to this prediction, a typical phenomenon in experimental asset markets is the emergence of price bubbles in which the price first rises substantially above the fundamental value, and "crashes" back to it in the final periods.

Several experimental studies use CP to shed more light on this phenomenon. In particular, Akiyama et al. (2017) compare asset markets populated by human subjects with asset markets in which a single subject interacts with CP and knows that they follow the equilibrium strategy. The authors find that mispricing drops by 50% in sessions with CP, which suggests that strategic uncertainty is an important driving factor of mispricing. Veiga and Vorsatz (2009, 2010) employ a CP following a simple algorithm to investigate whether experimental asset markets can be manipulated. They show that manipulation is possible, albeit only under certain information conditions. Moreover, the naive CP is sometimes able to make money. Jaworski and Kimbrough (2016) couple an asset market with a product market in which a monopolist firm posts prices and thus determines the dividends of the asset. The authors replicate the bubble phenomenon in this coupled market. Furthermore, differences between sessions where the firm is either controlled by a subject or a profit-maximizing CP are small, but CP facilitate learning.

All these studies show that asset market behavior becomes more rational in the presence of CP. Farjam and Kirchkamp (2018) go one step further and investigate how the mere expectation that CP could be present in the market affects behavior. They run three treatments, one with and two without CP. While subjects in one of the human-only treatments are informed about the treatment they are in, subjects in the other human-only treatment are left uncertain about their treatment, and they are not informed about the exact trading strategy of CP. Indeed, the authors find that price bubbles and volatility are smaller in the latter treatment in which subjects only expect CP to be present.

3.6 Auctions

Auctions are probably the most frequently investigated topic in experimental economics (see Kagel, 1995, Kagel and Levin, 2017). Typically, buyers submit bids for a given object (or set of objects) which is (are) acquired by the highest bidder(s). The main differences between auction formats lie in the pricing rule (usually, first- or second-price), the bidding opportunities (sealed-bid or dynamic), and the value of the object to the bidders (common or private). The most prominent experimental regularities include overbidding in first-price private value auctions, and the winner's curse in common value auctions.

Unsurprisingly, several studies employ CP in auctions to obtain a better understanding of the phenomena observed with human subjects, and in the field. I distinguish three strands of the literature. A first set of papers uses CP to eliminate strategic uncertainty and thus to investigate to which extent deviations from equilibrium predictions are caused by wrong expectations. Walker et al. (1987), Harrison (1989), and Ivanov et al. (2010) find little differences between treatments with and without CP indicating that (wrong) expectations are not the main driving force of disequilibrium behavior. Yet, Harrison (1989) shows that the foregone income is marginal, especially when subjects play against CP. Moreover, van den Bos et al. (2008) show that the absence of social context in sessions with CP reduces the winner's curse, and Chen and Takeuchi (2010) demonstrate that subjects imitate CP in a complex multi-object auction environment. Both studies suggest that CP do induce a shift in behavior towards rationality-based predictions.

A second strand of the literature employs CP to simplify the environment to subjects. Those papers investigate more complex or atypical auction formats (Cason, 1995, Kagel and Levin, 2001), or embed auctions in a more complex strategic environment (Davis et al., 2011, Brisset et al., 2015, Ausubel et al., 2017). A common finding is that the theory is able to explain some but not all features of the observed behavior. Only Cason (1995) investigates how behavior changes without CP, and finds that deviations are substantially larger.

Finally, Teubner et al. (2015) investigate how bidding against CP affects subjects' bidding behavior, and their emotions. CP replicated the bids of subjects in a previous session, but subjects only knew that they were interacting with CP (and not their strategy). The authors show that subjects who play against CP exhibit less arousal and a weaker relation between arousal and behavior, and submit lower bids.

3.7 Bargaining

(Camerer, 2003, p.151) calls bargaining "the most basic activity in economic life". Bargaining games usually involve a set of players (often, two) who need to negotiate the division of a given amount of resources. The outcome of such bargaining games heavily depends on the bargaining protocol, i.e. how offers can be made, and accepted or rejected, and how bargaining power is distributed among players. The few bargaining experiments which employ CP cover a rather wide range of bargaining protocols and aims pursued with the help of CP. This makes it difficult to draw common conclusions. I therefore focus on giving an impression of the breadth of applications of CP in bargaining.

Johnson et al. (2002) study an alternating-offer bargaining game and use CP to disentangle the impact of limited cognition and social preferences. Their CP implement the subgame-perfect equilibrium strategy, but subjects are only partially informed about the exact strategy. The authors find that CP bring subjects' behavior closer to equilibrium, but not fully.[8] Moreover, the interaction of CP has a lasting effect on behavior, as subjects keep playing closer to equilibrium when they are afterwards matched with other subjects.

Winter and Zamir (2005) use CP to manipulate beliefs in the ultimatum game (UG). Subjects first played the UG against other subjects for several rounds before CP implementing low, moderately low, or fair offers were introduced. This change in the strategic environment was not revealed to subjects in the main treatment, but in a robustness treatment included to check the impact of this unusual design choice. In both treatments, CP are able to sway behavior in their direction.

Embrey et al. (2015) test the reputation model of Abreu and Gul (2000) and employ CP to induce the behavioral types which are an important element of the model. In line with the predictions of the model, subjects mimic those types to build a reputation. In addition, complementary types emerge, whose initial demands acquiesce to the induced types.

A final paper deserves special emphasis: Manistersky et al. (2014) have subjects themselves program CP to play a bargaining game on their behalf. This is reminiscent of Axelrod's (1984) famous prisoner's dilemma tournament among game theorists. However, subjects in Manistersky et al. (2014) interacted repeatedly, and were allowed to change their CP across rounds. The authors find that designed CP improve across rounds, and approach the equilibrium of the game. The study points out a very interesting approach for future research. Already today, decision support systems are in use. In a world where many competitors employ AI, this may become essential. Experiments may then provide an understanding which factors drive the design, selection, and use of such systems.

3.8 Monetary Economics

Four papers I have found use CP to study the handling of money. Duffy (2001) investigates the validity of Kiyotaki and Wright's (1989) search model of money using both agent-based computations and experiments with human subjects. The major finding which emerges

[8] Similar findings are provided by de Melo and Gratch (2015) and Adam et al. (2018) under different bargaining protocols and for CP mimicking human behavior.

in this literature is a lack of speculation even when this is optimal: Both artificial and human players frequently do not accept a good which is more costly to store but also offers a better chance of trading in the future. Duffy investigates the driving forces of this stylized fact in two novel treatments. One of the two treatments has each subject interact with two CP who follow the equilibrium strategy. The aim is to facilitate learning. Indeed, the results suggest that subjects learn faster to adopt speculative strategies in this treatment. The main message of Duffy's paper is the usefulness of combining agent-based computations and laboratory experiments.

Fehr and Tyran (2001, 2007) and Petersen and Winn (2014) investigate the impact of money illusion, i.e. people erroneously taking nominal payoffs as a proxy for real payoffs. The authors analyze behavior in a simple pricing game with payoffs presented either in real or in nominal terms. As a main treatment variation, subjects interact either with other subjects, or with CP which follow the equilibrium strategy. This enables a separation of the direct effects of money illusion due to individual optimization mistakes from the indirect effects resulting from the expectation that others are prone to money illusion. The results show that money illusion causes substantial nominal inertia after a negative nominal shock, and coordination on inferior equilibria. However, the indirect effects of money illusion are much larger than its direct effects, suggesting that the results are mainly driven by strategic uncertainty.

3.9 Theory of Mind

Several studies use CP to understand subjects' strategic reasoning, or theory of mind. Those studies investigate behavior in simple, normal or extensive form games in which different levels of higher-order reasoning can easily be distinguished.

Costa-Gomes and Crawford (2006) conduct a treatment with CP in two-player guessing games to show that subjects can be induced to behave in line with certain behavioral types which employ different degrees of theory of mind. This provides the basis for their empirical identification of those types in sessions without CP.

Hanaki et al. (2016) consider 2×2 dominance-solvable coordination games. Their experimental design varies subjects' opponents (other subject or CP) and the monetary incentives to seek efficiency. In addition, they elicit subjects' cognitive ability. The results show that removing strategic uncertainty raises the efficiency of subjects' choices, but only partially and mainly for highly cognitively able subjects.

A series of papers in computer science also contributes to this literature.[9] In particular, Meijering et al. (2012) and Ghosh et al. (2016, 2017) study "marble-drop games"

[9]There are more papers in this literature (see in particular the references in Ghosh et al., 2017). However, it is not always clear, what subjects know about their opponents. I focus on studies in which I can rule out deception.

– representations of the centipede game in which a marble drops into a maze, and the players determine its path by alternatingly moving one of two trapdoors. The experimental results suggest that subjects apply forward rather than backward induction reasoning, and that considerations beyond game-theoretic reasoning (e.g. risk and social preferences and beliefs about them) are important for understanding subjects' behavior.

3.10 Other Topics

CP have been employed in experiments on various other topics without becoming an established methodological tool in these strands of the literature (yet). I here briefly summarize (sum of) those topics.

McCabe et al. (2001) and Wu et al. (2018) investigate behavior in the **trust game**. In particular, McCabe et al. (2001) show that subjects who cooperate with other subjects exhibit distinct brain activation and behavior when interacting with CP rather than humans, while this does not hold for non-cooperators. Wu et al. (2018) show that assigning an identity to CP (namely, an affiliation with a political party) may impact behavior.

Merlo and Schotter (1999, 2003) and Iyengar and Schotter (2008) use CP to investigate what and how subjects learn during repeated play of the **all-pay auction**. In a two player version, each subject is paired with a CP who implements the Nash equilibrium choice, and subjects are fully informed about this. Subjects first play 75 repetitions of the game. Afterwards, they are surprisingly faced with another round of the same game but with much higher stakes. This enables the authors to investigate exactly what subjects have learned. The results show that (i) subjects learn better if they are not paid while learning (Merlo and Schotter, 1999), (ii) observational learning may outperform learning-by-doing (Merlo and Schotter, 2003), and (iii) advice enhances the learning success (Iyengar and Schotter, 2008).

Grebe et al. (2008) and March and Ziegelmeyer (2018) use CP in **observational learning** settings. This literature follows the seminal model of Bikhchandani et al. (1992) in which a group of subjects are each endowed with private information about an unknown state of the world, and sequentially pick actions whose payoff depends on the state but not on the actions of others. The main theoretical prediction is the quick emergence of an *information cascade* in which all players neglect their private information and simply mimic their predecessors. In contrast, experiments typically find that subjects tend to rely too strongly on their private information, and fail to properly account for cascade behavior of other subjects. CP offer a way to disentangle competing explanations for these regularities. Accordingly, Grebe et al. (2008) use CP who follow the sequential equilibrium of the game, and explain them to subjects via a "counting rule". They find that subjects still follow private information too often, but less than in previous studies without CP. Moreover, elicited beliefs indicate that subjects also fail to recognize cascade behavior of CP. March and Ziegelmeyer (2018) investigate a setting with a richer information structure in which subjects may follow others both too little and too much. In different treatments, subjects either learn from other subjects, Bayes-rational CP (explained via a strategy table), or CP who always reveal their private information. The authors find that the reluctance to follow others disappears in settings with CP which suggests that this deviation is driven by strategic uncertainty. In contrast, subjects with very precise private information tend to follow other subjects and CP too often. Hence, non-Bayesian belief updating or misinferences cause excessive herding.

An interesting approach is pursued by Chen et al. (2018). They use CP to investigate **matching** in large markets. This involves two steps: The authors first demonstrate in a matching market of moderate size that subjects do not behave differently against other subjects than against CP programmed to mimic those subjects. Second, they investigate behavior in a large market in which each subjects is matched with 3,999 CP mimicking other subjects. The results show that subjects best respond more frequently in larger markets. Finally, the authors also show that rational behavior by CP enhances rational behavior by subjects.

Finally, four papers explicitly study how the interaction with CP changes behavior. Gogoll and Uhl (2016) show that subjects have a distaste for delegating decisions to CP. On the other hand, Hertz and Wiese (2019) find that subjects prefer to take advice from a CP for an analytical task, but not for a social task. Corgnet et al. (2019) show that subjects exert less effort when they are matched in a team which includes a CP, especially under team pay. This indicates that social incentives at the workplace may diminish when robots are employed. Finally, Kirchkamp and Strobel (2019) investigate joint decision-making of subjects when these are either matched with another subject or with a CP. They find little difference in behavior, and in the assignment of responsibility for the joint decision.[10]

4 Aims

Broadly speaking, CP are a tool to enhance experimental control, and therefore to increase the internal validity of experimental studies. For example, Messick (1967) terms experiments on games in which subjects interact simultaneously a "*noncontrol* method" and argues that CP are "a method by which means the experimenter can directly manipulate the nature of the dependence of one of the participants" (p. 34). Similarly,

[10]CP have also been employed (i) to induce behavioral types in experimental tests of reputation models (Bolton and Ockenfels, 2005, Grosskopf and Sarin, 2010), (ii) to investigate the impact of social status on learning in a coordination game (Eckel and Wilson, 2007), (iii) to isolate the impact of strategic uncertainty on subjects' ability to play a correlated equilibrium (Cason and Sharma, 2007), and to study the impact of social preferences in laboratory contests (Cox et al., 2017). I do not review these papers in greater detail.

Roth and Murnighan (1978) use CP "to control for differences in subjects' behavior due to differences in their opponents". The way in which CP enhance control depends on their exact implementation, and therefore varies across studies. Below, I classify the aims that have been pursued in the literature using CP into five categories. The appendix contains an assignment of the various studies to those categories. The categories are not fully distinct, and several studies might be classified differently, or assigned to multiple categories.

As usual, enhancing internal validity via CP comes at the expense of a decrease in external validity. I discuss the external validity of experiments with CP in section 6.

4.1 Reducing Noise

A first set of studies use CP mainly to reduce the noise in the experiment by eliminating one serious confound, namely (changes in) the behavior of one group of players. There are different reasons for doing so.

First, using CP to reduce noise enables the experimenters a better understanding of their results. A good example are market experiments with CP. For example, Cason and Williams (1990, p.337) "explicitly controlled for strategic buyer behavior [...] to avoid 'gaming' by the buyers which could complicate the interpretation of the experimental results". Accordingly, CP facilitate the identification of experimental phenomena underlying noisy data (see Guala, 2005, p.45).

Second, CP also facilitate the testing of theory as a means of controlling auxiliary assumptions (see e.g. Guala, 2005, p.60). For instance, Isaac and Smith (1985) argue that CP are needed in their experiments "because virtually all of the hypotheses regarding predatory pricing explicitly or implicitly assume that buyers act to fully reveal market demand." In this case, buyer behavior is not the focus of the study, but acts as an auxiliary assumption in the theory of seller behavior. More generally, CP are a mean of control in the experimental investigation of *individual* strategic behavior. Indeed, game theoretic models, especially equilibrium models, often make predictions about a group of players rather than individual players. Testing whether or not those theories predict the behavior of individual subjects requires their opponents to stick to the theory. Hence, opponent behavior becomes an auxiliary assumption which can be controlled via CP. A good example are the experiments on mixed-strategy Nash equilibrium.

Third, CP simplify the strategic environment for subjects. For example. Kagel and Levin (2001, p.415) use CP to generate "an environment free from the strategic uncertainties inherent in interactions between human bidders". Similarly, Davis et al. (1993, p.308) argue that "any less-than-fully-revealing buyer behavior would only complicate an already complex price adjustment problem for posted-offer sellers". An indirect version of this argument posits that CP, by generating a stable environment, facilitate learning for subjects (see e.g. Cason and Friedman, 1997, Duffy, 2001).

Finally, there are secondary benefits. In particular, CP may help to reduce subject payment expenses (Cason and Williams, 1990), simplify the instructions (Morgan et al., 2006), and increase the statistical power of tests.

A common feature of studies employing CP merely to reduce noise is the reliance on CP following some equilibrium strategy (25 out of 31 studies). In addition, most studies inform subjects fully (19 out of 31) or at least partially (5 out of 31) about the strategy of CP, and do not provide an explicit comparison of treatments with and without CP (19 out of 31).

4.2 Inducing Types

A second important purpose of using CP is the induction of specific behavioral types. This is for instance necessary, if theories are tested which explicitly assume the existence of those types, and we cannot expect human subjects to represent the types (exactly). Examples are experimental tests of bargaining models (e.g. Embrey et al., 2015), reputation models (e.g. Grosskopf and Sarin, 2010), and models from behavioral economics (e.g. Kalayci, 2015).

Regardless of testing a specific model, we might simply be interested in the way subjects respond to a specific strategy, or type of behavior. This can be motivated by observations from previous experiments, which suggest that the behavior is present or even common in the lab or the field. It can also be motivated by theoretical considerations in which a model is (partly) justified on the grounds that it specifies a good response to human behavior. A good example are various studies on mixed-strategy Nash equilibrium which pit subjects against fixed mixed strategies to see whether they best respond, or against CP programmed to implement a learning model to understand if and how the model is able to outperform humans.

A third reason for the use of CP to induce types is the manipulation of beliefs. For instance, Andreoni and Miller (1993) (respectively, Guillen et al., 2010) introduce the possibility that subjects face CP implementing a tit-for-tat (resp. grim-trigger) strategy in a prisoner's dilemma (resp. public good game). Both studies find that the mere possibility of the existence of those types is sufficient to induce higher cooperation.

Finally, Chen et al. (2018) use CP to enlarge the size of the market in an experimental test of matching mechanisms. The authors first demonstrate that subjects do not behave differently when playing against 39 other subjects, or 39 CP who mimic those subjects by drawing their strategies from the empirical distribution of their strategies. In a subsequent treatment, they match each subject with 3,999 CP that draw their strategies from the same empirical distribution. The results show that subjects are more inclined to best respond in larger markets, regardless of the matching mechanism.

Of the 22 studies employing CP to induce types (in at least one treatment), 19 program CP to implement a fixed strategy, or an adaptive algorithm able to adjust to the behavior of subjects (like fictitious play or tit-for-tat). The majority of the studies inform the subjects at most partially about the exact algorithm underlying CP (14 out of 22), and do not provide a comparison between treatments with and without CP (14 out of 22).

4.3 Removing Strategic Uncertainty

Starting with Walker et al. (1987), a series of papers have used CP to explicitly isolate the impact of strategic uncertainty on decisions. The decisive feature of those studies is the comparison of one treatment in which subjects interact with other subjects and strategic uncertainty is present with a second treatment in which subjects interact with CP and strategic uncertainty is absent.[11]

The results from this literature show that the impact of strategic uncertainty is heavily context-dependent. While strategic uncertainty accounts for a substantial part of the deviations from equilibrium predictions in money illusion settings (Fehr and Tyran, 2001, 2007), oligopoly pricing (Kalayci and Potter, 2011, Dua et al., 2013), and experimental asset markets (Akiyama et al., 2017), the impact of strategic uncertainty is more nuanced in social learning (March and Ziegelmeyer, 2018), and largely absent in auction settings (e.g. Walker et al., 1987, Ivanov et al., 2010) and strictly competitive normal-form games (Levitt et al., 2010).

The majority of studies employing CP to remove strategic uncertainty match subjects with CP implementing the equilibrium strategy in the respective setting (14 out of 18 studies), and all but one study inform subjects fully about the algorithm underlying CP.

4.4 Excluding Social Preferences

Houser and Kurzban (2002) and Johnson et al. (2002) introduced a novel use of CP to the experimental literature: the removal of social preferences from the considerations of human subjects. The argument is twofold. First, since the strategy of CP is predetermined, reciprocity should not affect subjects' choices as it does not pay to reward good or punish bad behavior by CP. Second, as CP do not receive money, subjects should feel neither envious nor guilty towards CP which is why altruism (or spite) should not affect decisions either. Both studies find that subjects' decisions move closer to the selfish equilibrium prediction, but not fully. Accordingly, social preferences influence subjects' decisions in the absence of CP, but other motives not taken into account in the equilibrium models also affect behavior. Therefore, besides elucidating the concrete impact of social preferences,

[11] Only Grebe et al. (2008) do not perform such a comparison. However, they compare their results to a large experimental literature with well-established experimental regularities.

CP also serve to create a strategic environment free from the social preference confound in which to study those other motives.

The literature following these two initial studies has largely followed Houser and Kurzban (2002) and focused on the public good game. Most of these studies have CP mimic human subjects' decisions from previous sessions, inform subjects about this fact (and sometimes even about the exact choices by CP), and perform explicit comparisons between treatments with and without CP. The results by and large confirm that social preferences partially but not fully explain positive contributions, though the exact impact is still debated. Interestingly, the few studies which employ CP to exclude social preferences in other settings, do so in settings in which negatively annotated other-regarding preferences (e.g. spite, willingness to beat others) are surmised to affect behavior. Whereas van den Bos et al. (2008) confirm this surmission in first-price common-value auctions, Cox (2017) finds no such impact on overbidding in the Tullock contest.

One final note is in order: The claim that CP remove outcome-based social preferences like altruism or spite as CP "do not receive money" assumes away the possibility of altruism towards the experimenter, as CP in some sense act on her behalf. This has largely been the case in experimental studies of social preferences (with or without CP). However, altruism towards the individual deploying CP might become an important aspect of real-world human-machine-interaction, if artificial agents are increasingly explored to act on behalf of individuals. Few studies explicitly account for this possibility by having subjects receive the earnings of CP. Yamakawa and Okano (2016) find that contributions to a public good game increase, if some subjects receive the money of a CP. This suggests that taking into account who benefits from the earnings of CP could be an important aspect of future research.

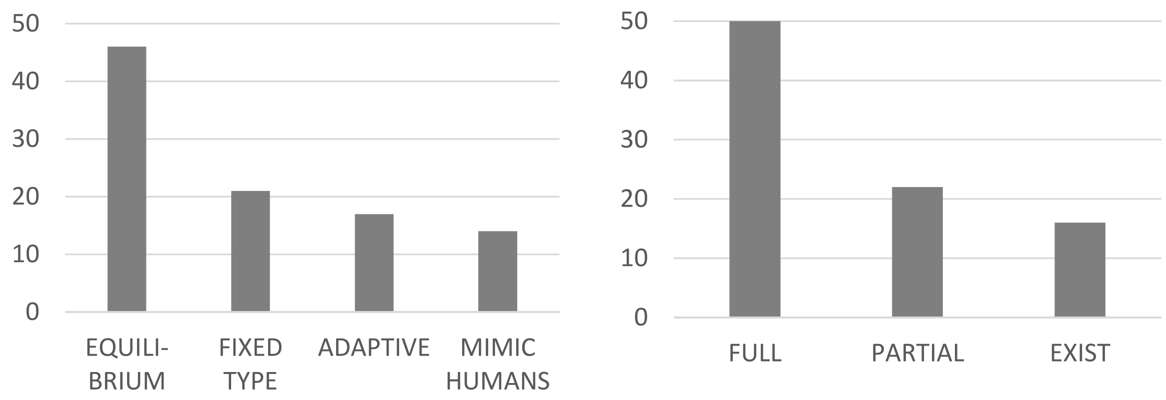
4.5 Human-Computer Interaction

Recently, experimental economists have joined computer scientists in investigating explicitly how human strategic behavior changes when they face CP instead of or in addition to other human subjects. There are two general research questions that have been pursued. A first strand of the literature uses algorithms from computer science (or economics) to implement CP that could well be encountered in the real world. The goal here is an understanding of how subjects compete against such sophisticated behavior, and an answer to the question who outperforms whom. The information given to subjects is often minimal, and an explicit comparison to treatments without CP is often not performed. The common finding is that subjects are able to adapt to and partly exploit the algorithms, but that they are still sometimes outperformed if the algorithm of CP is sufficiently sophisticated.

A second set of studies has CP mimic human choice behavior (usually, by drawing

FIGURE 2
Methodological Aspects of Computer Players

(a) Programmed Strategies (b) Information to Subjects



from choices in previous sessions without CP) to understand how the mere change in the nature of a subject's opponents affects her behavior. These studies provide slightly more information on CP to subjects, and usually compare behavior across treatments with and without CP. In general, behavior does change, and subjects behave more rational, i.e. closer to equilibrium predictions.

5 Methods

I here summarize the different variants in which CP have been implemented and explained to subjects, and whether and how the impact of CP has been explicitly investigated. Obviously, the methodological choices heavily depend on both the topics under investigation, and the aim that is pursued with the help of CP. I try to clarify those dependencies.

5.1 Algorithms

Figure 2a gives an overview over the different strategies or algorithms that have been used to program CP. Notice that several studies have implemented multiple treatments with different variants of CP and are therefore counted multiple times. Moreover, two studies do not reveal the algorithm of their CP.

Around half of the studies I have collected implement CP (in at least one treatment) which follow some kind of **equilibrium strategy**. The exact strategy depends on the strategic environment, and the equilibrium concept which is employed. Strategies thus differ widely in this category, ranging from minimax or dominant strategies (e.g. buyers in posted-offer markets) to Perfect Bayesian equilibrium strategies (e.g. auctions or observational learning settings) which require additional assumptions regarding for example risk preferences and expectations about others' preferences. Equilibrium CP are employed across topics, and aims pursued with CP. In addition, subjects are often informed fully about the respective equilibrium strategy (37 out of the 46 treatments; see below), and behavior in the presence of CP is often compared to behavior in a treatment without CP (30 out of the 46 treatments).

A second class of CP follow a predetermined, **fixed type**. Again, the exact nature depends on the strategic environment, and frequently also the underlying theory. The class comprises CP who draw their actions randomly from a given domain and distribution, e.g. CP who follow a (non-equilibrium) mixed strategy in normal-form games. The class also comprises CP who implement a theoretically assumed type, e.g. strong competitors in a reputation setting (Bolton and Ockenfels, 2005, Grosskopf and Sarin, 2010), or bargaining types whom subjects are predicted to mimic (Embrey et al., 2015). Finally, some studies use CP to implement some kind of (non-equilibrium) benchmark behavior, e.g. truthful bidding in auctions (Cason and Friedman, 1997). Unsurprisingly, fixed type CP are often employed to reduce noise or induce types in the strategic environment. Subjects are only informed about the exact type of CP in half of the cases, and a comparison of treatments with and without CP is about equally common.

The third class of CP employs **adaptive algorithms** that are able to be responsive to the behavior of subjects. This class comprises repeated game strategies, such as tit-for-tat or grim-trigger. It also comprises learning models like fictitious play and sophisticated algorithms developed (mainly) in the computer science literature. CP from this class are mainly employed in studies of, respectively, social dilemmas and mixed-strategy Nash equilibrium, and with the aim of studying subjects' responses to those adaptive types of behavior. Subjects are rarely informed in detail about the nature of the underlying algorithms, and treatment comparisons are also infrequent in these studies.

A fourth class follows the method of Houser and Kurzban (2002) and has CP **mimic the behavior of human subjects** in a previous experiment (without CP). The implementation differs from study to study. While some studies draw a full sequence of choices made by a single human subject for each CP, others program each CP to implement the average choice, or to draw in each repetition from the empirical distribution of choices. While employed initially to exclude social preferences in social dilemmas, the method has more recently been employed in other strategic environments, and for different purposes. All of those studies explicitly compare treatments with and without CP. Furthermore, most studies inform subjects that CP mimic other humans, but they do not inform subjects about their exact strategy or the distribution it is drawn from. Incidentally, this does not hold for Houser and Kurzban (2002), who tell subjects in each round before they make their choices what the CP in their group are going to contribute to the public good. The problem with such an approach is that it changes the strategic interaction to a sequential game which confounds the comparison of treatments.

Finally, CP in Manistersky et al. (2014) are designed by subjects, and the design choices are the main object of the study (see above).

5.2 Information to Subjects

Figure 2b illustrates the type of information that has been given to subjects about CP. The "no-deception-condition" requires that subjects are at least made aware of their interaction with CP. Two studies are included despite not satisfying this requirement. Shachat and Swarthout (2004) inform subjects that they would play against a "decision maker" without specifying its nature. Hence, subjects are at least not actively misled to believe that they are interacting with another subject. Winter and Zamir's (2005) main treatment disguises the fact that CP are used, but they also include a robustness check in which the presence of CP is revealed and find little difference.

As can be seen, the overwhelming majority of studies informs subjects about the behavior of CP in great detail. This guarantees the greatest degree of experimental control, but at the expense of a loss in external validity (see the discussion). There are two difficulties involved: The first difficulty stems from the aspiration to make the subjects fully understand the strategy of CP. This heavily depends on the strategic environment. Whereas the revelation of the strategy of CP is easy in simple games (e.g. normal-form games) and if it comes natural to the subjects (e.g. CP which buy any profitable unit in posted-offer markets), strategies deriving from sophisticated theoretical concepts in complex environments are less easily explained. Therefore, several studies go to some lengths to clarify the strategy of CP, e.g. via tables (e.g. Fehr and Tyran, 2001, March and Ziegelmeyer, 2018), figures (e.g. Cason and Friedman, 1997), and/or interactive tools (e.g. Ausubel et al., 2017). The second difficulty derives from the possibility that subjects interpret the strategy of CP as an optimal way to play the game, and simply mimic it. This raises problems of disentangling changes in subjects' behavior due to the change in opponents from changes due to the information given to subjects in the instructions. Potential solutions are the assignment of CP and subjects to different roles (if possible), or explaining the strategy of CP without reference to the underlying algorithm (or thought process).

A second set of studies tell subjects that they interact with CP, and give some information about the principles underlying the CP algorithm, but do not provide the complete strategy. Subjects are e.g. told that CP are programmed to "make as much money as possible for itself" (Johnson et al., 2002, Veiga and Vorsatz, 2009), to "play like a human" (West and Lebiere, 2001), or to "implement the theoretically correct strategy" (Levitt et al., 2010). Subjects are often also informed, that the choices of CP "have been previously determined" and are therefore "independent of [subjects'] decisions" (Di Mauro and Castro, 2011). The exact information depends on the strategic environment, and the purpose and nature of CP. For example, partial information is commonly employed, if CP mimic the decisions of human subjects in a previous session. The reference to other subjects is intended to evoke similar beliefs in subjects as in human-only sessions, while full information can hardly be revealed without changing the nature of the strategic interaction (see Section 5.1).

Finally, some studies do not move beyond informing subjects about the existence of CP. Half of those investigate rather sophisticated adaptive algorithms, which are not easily explained to subjects (e.g. Das et al., 2001, Spiliopoulos, 2013). A second reason for the provision of minimal information is that subjects' beliefs about the behavior of CP are part of the research question (e.g. Eckel and Wilson, 2007, Teubner et al., 2015).

5.3 Treatment Comparisons

54 out of the 90 studies reviewed here compare the strategic behavior of subjects when interacting with CP to the behavior absent CP. Clearly, this design provides the cleanest identification of the impact of CP, as the patterns of behavior observed in the literature are first replicated for the subject pool and the concrete experimental implementation at hand (see e.g. Guala, 2005). Whether or not such a comparison is necessary depends on the research question. Thus, treatment comparisons are especially conducted, if CP are used to isolate the impact of e.g. strategic uncertainty or social preferences on behavior.

Four studies are worth mentioning. Cason and Datta (2006) and Gong et al. (2013) have published the treatments without CP in separate studies (Cason and Mago, 2010 and Gong et al., 2009, respectively). Duffy and Xie (2016) vary the number of CP across different treatments, but they do not study the case without CP. Finally, Farjam and Kirchkamp (2018) mainly investigate how the mere possibility that CP are present changes behavior (see Section 3).

6 Discussion

In the past two decades, computer players have become an ubiquitous tool in experimental economics. Though they have mainly pursued methodological aims, studies employing computer players deliver initial insights into the change in strategic behavior in the presence of computer players. First, subjects generically behave more selfish and more rational when interacting with computers. Second, subjects can often learn to exploit even rather sophisticated computer players, and with only little prior information. Third, there are limits to adaption and exploitation, and subjects are sometimes outperformed by computer players.

So far, computer players have largely been implemented with a focus on increasing experimental control, and testing economic theories. In particular, the majority of studies use computer players who follow an equilibrium strategy, fully inform subjects about the behavior of computer players, and explicitly provide a comparison between treatments with and without computer players. This is not without merit for the study of strategic interaction with artificial intelligence. For example, *Alpha Go Zero* – a version of the software *Alpha Go* which beat a world champion in the game of Go – learned to play the game only from plays against itself by employing a reinforcement learning algorithm, and eventually outperformed all previous versions (Silver et al., 2017). It is well known that the strategies of players who all use reinforcement learning converge to equilibrium in certain games (see e.g. Beggs, 2005). Accordingly, the behavior of subjects vis-à-vis computer players who employ equilibrium strategies might inform us about the behavior against AI agents trained through self-play.

On the other hand, many real-world interactions with artificial intelligence will not coincide with such carefully controlled laboratory conditions. The current debate on AI revolves particularly around the difficulty to understand, trace, and explain the way AI reaches its decisions (see e.g. Lapuschkin et al., 2019). Accordingly, full information about the algorithm of computer players will unlikely apply in the real world. Yet, currently developed AI guidelines require that "it should be possible to demand a suitable explanation of the AI system's decision-making process" and that the "explanation should be [...] adapted to the expertise of the stakeholder concerned" (European Commission, 2019). Experimental and behavioral economists surely have a lot to contribute to this debate, e.g. by investigating how individuals' beliefs about the strategy of an AI system depend on the way the information about the AI system's decision-making process is presented.

At the same time, it is commonly required that humans always be informed about their interaction with AI (see e.g. OECD, 2019a, European Commission, 2019). This puts additional doubt on studies which do not inform subjects about their interaction with computer players.

As the public debate on the requirements on AI picks up pace, several avenues for future research open up for experimental economists. By demonstrating how the interaction with AI changes human strategic behavior and incentives, they may weigh in on the debate about the Dos and Dont's in the development and deployment of AI. Moreover, they may show whether and when humans accept advice from computer players,[12] or are willing to delegate their decisions fully to AI. Computer players are thus set to become more than a methodological tool.

[12]See Schotter (2003) for a survey on advice taking in experiments with and without computer players.

Appendix. Overview of the Studies

Table 1 provides an overview of all papers reviewed in this survey. I sort the papers by topic and year. Furthermore, I provide information on the following aspects of the studies:

(i) *Purpose of CP*: Based on the main goal for using computer players stated in the paper, I assign one the following five categories to each study: Reduce Noise (subsection 4.1), Induce Types (subsection 4.2), Remove Strategic Uncertainty (subsection 4.3), Exclude Social Preferences (subsection 4.4), or Human-Computer Interaction (subsection 4.5)

(ii) *Strategy of CP*: The third column collects how computer players are programmed to play. In line with the argumentation in section 5.1, I distinguish five categories: Equilibrium, Adaptive, Fixed Type, Mimic Humans, Designed by Subjects.

(iii) *Info to Subjects*: The fourth column summarizes how subjects are informed about the strategy of computer players. Following section 5.2, I distinguish four categories: FULL, PARTIAL, EXIST, or NONE.

(iv) *Comparison (Comp.)*: The final column indicates, whether or not the study provides a direct comparison of behavior with and without CP. Entries marked with an asterisk indicate specific studies discussed in subsection 5.3

Clearly, the different categories are often not sharply distinguished, and several studies could well be classified differently. Indeed, some studies are assigned to multiple categories, either because they employ multiple treatments, or because a distinction is nearly impossible. Furthermore, I did not manage to find the necessary information for all studies. Missing entries are marked by “n/a”.

TABLE 1
Overview of Studies

Study	Purpose of CP	Strategy of CP	Info to Subjects	Comp.
A. Mixed-Strategy Nash Equilibrium				
Messick (1967)	Reduce Noise/ Induce Types	Equilibrium / Adaptive	PARTIAL	No
Fox (1972)	Reduce Noise/ Induce Types	Equilibrium / Fixed Type	PARTIAL	No
West and Lebiere (2001)	Induce Types	Adaptive	PARTIAL	No
Shachat and Swarthout (2004)	Induce Types	Fixed Type	NONE	No
Coricelli (2005)	Induce Types	Adaptive	EXIST	No
Cohen and Ranganath (2007)	Induce Types	Adaptive	n/a	No
Levitt et al. (2010)	Remove Strategic Uncertainty	Adaptive / Fixed Type	PARTIAL	Yes
Spiliopoulos (2013)	Induce Types	Adaptive	EXIST	No
Moisan and Gonzales (2017)	Reduce Noise/ Induce Types	Equilibrium / Fixed Type / Adaptive	EXIST	No
Griessinger et al. (2018)	Induce Types	Adaptive	EXIST	Yes
B. Prisoner's Dilemma				
Shubik et al. (1974)	Reduce Noise	Adaptive	EXIST	No
Roth and Murnighan (1978)	Reduce Noise	Adaptive	EXIST	No
Andreoni and Miller (1993)	Induce Types	Adaptive	FULL	Yes
Gong et al. (2013)	Exclude Social Preferences	Mimic Humans	FULL	Yes*
Duffy and Xie (2016)	Remove Strategic Uncertainty	Adaptive	FULL	Yes*
C. Public Good Games				
Houser and Kurzban (2002)	Exclude Social Preferences	Mimic Humans	FULL	Yes
Ferraro et al. (2003)	Exclude Social Preferences	Mimic Humans	PARTIAL	Yes

Continued on next page

TABLE 1
Overview of Studies (ctd.)

Study	Purpose of CP	Strategy of CP	Info to Subjects	Comp.
Ferraro and Vossler (2010)	Exclude Social Preferences	Mimic Humans	PARTIAL	Yes
Guillen et al. (2010)	Induce Types	Adaptive	PARTIAL	Yes
Di Mauro and Castro (2011)	Exclude Social Preferences	Fixed Type	PARTIAL	No
Burton-Chellew et al. (2016)	Exclude Social Preferences	Fixed Type	FULL	Yes
Yamakawa and Okano (2016)	Exclude Social Preferences	Mimic Humans	FULL	Yes
D. Market Experiments				
Wolf and Shubik (1978)	Induce Types	Adaptive	EXIST	No
Coursey et al. (1984)	Reduce Noise	Equilibrium	FULL	Yes
Harrison and McKee (1985)	Reduce Noise	Equilibrium	FULL	No
Isaac and Smith (1985)	Reduce Noise	Equilibrium	FULL	No
Davis and Williams (1986)	Reduce Noise	Equilibrium	FULL	No
Cason and Williams (1990)	Reduce Noise	Equilibrium	n/a	Yes
Brown Kruse (1991)	Reduce Noise	Equilibrium	FULL	Yes
Davis and Williams (1991)	Reduce Noise	Equilibrium	EXIST	Yes
Davis et al. (1993)	Reduce Noise	Equilibrium	n/a	No
Cason and Friedman (1997)	Reduce Noise	Equilibrium / Fixed Type	FULL	Yes
Das et al. (2001)	Human-Computer Interaction	Adaptive	EXIST	No
Cason and Friedman (2002)	Reduce Noise	Equilibrium	FULL	Yes
Cason and Datta (2006)	Reduce Noise	Equilibrium	FULL	Yes*
Morgan et al. (2006)	Reduce Noise	Equilibrium	FULL	No
Duersch et al. (2010)	Human-Computer Interaction	Adaptive	PARTIAL	No
Mago (2010)	Reduce Noise	Equilibrium	FULL	Yes
De Luca and Cliff (2011)	Human-Computer Interaction	Adaptive	n/a	No

Continued on next page

TABLE 1
Overview of Studies (ctd.)

Study	Purpose of CP	Strategy of CP	Info to Subjects	Comp.
Kalayci and Potter (2011)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
Dua et al. (2013)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
Kalayci (2015)	Induce Types	Equilibrium / Fixed Type	FULL	Yes
E. Experimental Asset Markets				
Veiga and Vorsatz (2009)	Induce Types	Fixed Type	PARTIAL	Yes
Veiga and Vorsatz (2010)	Induce Types	Fixed Type	PARTIAL	Yes
Jaworski and Kimbrough (2016)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
Akiyama et al. (2017)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
Farjam and Kirchkamp (2018)	Human-Computer Interaction	Equilibrium	PARTIAL	Yes*
F. Auctions				
Walker et al. (1987)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
Harrison (1989)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
Cason (1995)	Reduce Noise	Equilibrium	FULL	Yes
Kagel and Levin (2001)	Reduce Noise	Equilibrium	FULL	No
van den Bos et al. (2008)	Exclude Social Preferences	Equilibrium / Mimic Humans	FULL / PARTIAL	Yes
Chen and Takeuchi (2010)	Remove Strategic Uncertainty	Equilibrium / Fixed Type	FULL / EXIST	Yes
Ivanov et al. (2010)	Remove Strategic Uncertainty	Mimic Humans	FULL	Yes
Davis et al. (2011)	Reduce Noise	Equilibrium	FULL	No
Brisset et al. (2015)	Reduce Noise	Equilibrium	FULL	No
Teubner et al. (2015)	Human-Computer Interaction	Mimic Humans	EXIST	Yes
Ausubel et al. (2017)	Reduce Noise	Equilibrium	FULL	Yes

Continued on next page

TABLE 1
Overview of Studies (ctd.)

Study	Purpose of CP	Strategy of CP	Info to Subjects	Comp.
G. Bargaining				
Johnson et al. (2002)	Exclude Social Preferences	Equilibrium	PARTIAL	Yes
Winter and Zamir (2005)	Reduce Noise	Fixed Type	EXIST / NONE	Yes
Manistersky et al. (2014)	Human-Computer Interaction	Designed by Subjects	FULL	No
de Melo and Gratch (2015)	Human-Computer Interaction	n/a	EXIST	Yes
Embrey et al. (2015)	Induce Types	Fixed Type	FULL	No
Adam et al. (2018)	Human-Computer Interaction	Mimic Humans	EXIST	Yes
H. Monetary Economics				
Duffy (2001)	Reduce Noise	Equilibrium	FULL	Yes
Fehr and Tyran (2001)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
Fehr and Tyran (2007)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
Petersen and Winn (2014)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
I. Strategic Reasoning				
Costa-Gomes and Crawford (2006)	Induce Types	Fixed Type	FULL	No
Meijering et al. (2012)	Induce Type	Equilibrium	FULL	No
Ghosh et al. (2016)	Reduce Noise	Fixed Type	PARTIAL	No
Hanaki et al. (2016)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
Ghosh et al. (2017)	Reduce Noise	Fixed Type	PARTIAL	No
J. Other Studies				
Merlo and Schotter (1999)	Reduce Noise	Equilibrium	FULL	No
McCabe et al. (2001)	Remove Strategic Uncertainty	Fixed Type	FULL	Yes
Merlo and Schotter (2003)	Reduce Noise	Equilibrium	FULL	No

Continued on next page

TABLE 1
OVERVIEW OF STUDIES (CTD.)

Study	Purpose of CP	Strategy of CP	Info to Subjects	Comp.
Bolton and Ockenfels (2005)	Induce Types	Fixed Type	FULL	No
Cason and Sharma (2007)	Remove Strategic Uncertainty	Equilibrium	FULL	Yes
Eckel and Wilson (2007)	Induce Types	Equilibrium	EXIST	Yes
Grebe et al. (2008)	Remove Strategic Uncertainty	Equilibrium	FULL	No
Iyengar and Schotter (2008)	Reduce Noise	Equilibrium	FULL	No
Grosskopf and Sarin (2010)	Induce Types	Fixed Type	FULL	No
Gogoll and Uhl (2016)	Human-Computer Interaction	Mimic Humans	PARTIAL	Yes
Cox (2017)	Exclude Social Preferences	Mimic Humans	PARTIAL	Yes
Chen et al. (2018)	Induce Types	Equilibrium / Mimic Humans	FULL / PARTIAL	Yes
March and Ziegelmeyer (2018)	Remove Strategic Uncertainty	Equilibrium / Fixed Type	FULL	Yes
Wu et al. (2018)	Reduce Noise	Fixed Type	PARTIAL	No
Corgnet et al. (2019)	Human-Computer Interaction	Mimic Humans	PARTIAL	Yes
Hertz and Wiese (2019)	Human-Computer Interaction	n/a	EXIST	Yes
Kirchkamp and Strobel (2019)	Human-Computer Interaction	Mimic Humans	PARTIAL	Yes

References

Abreu, D. and F. Gul (2000): "Bargaining and Reputation," *Econometrica*, 68, 85–117.

Acemoglu, D. and P. Restrepo (2019): "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, forthcoming.

Adam, M. T., T. Teubner, and H. Gimpel (2018): "No Rage Against the Machine: How Computer Agents Mitigate Human Emotional Processes in Electronic Negotiations," *Group Decision and Negotiation*, 27, 543–571.

Agrawal, A., J. Gans, and A. Goldfarb (2019): *The Economics of Artificial Intelligence: An Agenda*, The University of Chicago Press, Chicago and London.

Akiyama, E., N. Hanaki, and R. Ishikawa (2017): "It is Not Just Confusion! Strategic Uncertainty in an Experimental Asset Market," *The Economic Journal*, 127, 563–580.

Anderhub, V., D. Engelmann, and W. Güth (2002): "An experimental study of the repeated trust game with incomplete information," *Journal of Economic Behavior and Organization*, 48, 197–216.

Andreoni, J. (1995): "Cooperation in Public-Goods Experiments: Kindness or Confusion?" *American Economic Review*, 85, 891–904.

Andreoni, J. and J. H. Miller (1993): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," *The Economic Journal*, 103, 580–585.

Andrews, D. G. N. and C. Timiliotis (2018): "Digital technology diffusion: A matter of capabilities, incentives or both?" Oecd economics department working papers, no. 1476, OECD Publishing, Paris.

Ausubel, L. M., J. E. Burkett, and E. Filiz-Ozbay (2017): "An experiment on auctions with endogenous budget constraints," *Experimental Economics*, 20, 973–1006.

Axelrod, R. (1984): *The Evolution of Cooperation*, Basic Books, New York.

Beggs, A. W. (2005): "On the convergence of reinforcement learning," *Journal of Economic Theory*, 122, 1–36.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, 992–1026.

Bolton, G. E. and A. Ockenfels (2005): "The Influence of Information Externalities on the Value of Reputation Building - An Experiment," Working paper series in economics, University of Cologne.

Brisset, K., F. Cochard, and F. Marèchal (2015): "Is the newcomer more aggressive when the incumbent is granted a Right-of-First-Refusal in a procurement auction? Experimental Evidence," *Theory and Decision*, 78, 639–665.

Brogaard, J., T. Hendershott, and R. Riordan (2014): "High-Frequency Trading and Price Discovery," *The Review of Financial Studies*, 27, 2267–2306.

Brown Kruse, J. (1991): "Contestability in the Presence of an Alternate Market: An Experimental Examination," *The RAND Journal of Economics*, 22, 136–147.

——— (2008): "Simulated and Real Buyers in Posted Offer Markets," in *Handbook of Experimental Economics Results*, ed. by C. R. Plott and V. L. Smith, North-Holland, Amsterdam, vol. 1, chap. 8, 71–76.

Burton-Chellew, M. N., C. El Mouden, and S. A. West (2016): "Conditional cooperation and confusion in public-goods experiments," *Proceedings of the National Academy of Sciences*, 113, 1291–1296.

Camerer, C. F. (2003): *Behavioral Game Theory*, Princeton University Press, Princeton.

——— (2019): "Artificial Intelligence and Behavioral Economics," in *The Economics of Artificial Intelligence: An Agenda*, ed. by A. Agrawal, J. Gans, and A. Goldfarb, The University of Chicago Press, Chicago and London, 587–608.

Cason, T. N. (1995): "An Experimental Investigation of the Seller Incentives in the EPA's Emission Trading Auction," *American Economic Review*, 85, 905–922.

Cason, T. N. and S. Datta (2006): "An experimental study of price dispersion in an optimal search model with advertising," *International Journal of Industrial Organization*, 24, 639–665.

Cason, T. N. and D. Friedman (1997): "Price Formation in Single Call Markets," *Econometrica*, 65, 311–345.

——— (2002): "Buyer search and price dispersion: a laboratory study," *Journal of Economic Theory*, 104, 16–47.

Cason, T. N. and S. D. Mago (2010): "Costly Buyer Search in Laboratory Markets with Seller Advertising," *The Journal of Industrial Economics*, 58, 424–449.

Cason, T. N. and T. Sharma (2007): "Recommended play and correlated equilibria: an experimental study," *Economic Theory*, 33, 11–27.

Cason, T. N. and A. W. Williams (1990): "Competitive Equilibrium Convergence in a Posted-Offer Market with Extreme Earnings Inequities," *Journal of Economic Behavior and Organization*, 14, 331–352.

Charness, G. and D. Levin (2009): "The Origin of the Winner's Curse: A Laboratory Study," *American Economic Journal: Microeconomics*, 1, 207–236.

Chen, Y., M. Jiang, O. Kesten, S. Robin, and M. Zhu (2018): "Matching in the large: An experimental study," *Games and Economic Behavior*, 110, 295–317.

Chen, Y. and K. Takeuchi (2010): "Multi-object auctions with package bidding: An experimental comparison of Vickrey and iBEA," *Games and Economic Behavior*, 68, 557–579.

Cohen, M. X. and C. Ranganath (2007): "Reinforcement Learning Signals Predict Future Decisions," *The Journal of Neuroscience*, 27, 371–378.

Corgnet, B., R. Hernàn-Gonzalez, and R. Mateo (2019): "Rac(g)e Against the Machine? Social Incentives When Humans Meet Robots," Working paper halshs-01994021, HAL.

Coricelli, G. (2005): "Strategic Interaction in Iterated Zero-Sum Games," Mimeo.

Costa-Gomes, M. A. and V. P. Crawford (2006): "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, 96, 1737–1768.

Coursey, D., R. M. Isaac, M. Luke, and V. L. Smith (1984): "Market Contestability in the Presence of Sunk (Entry) Costs," *The RAND Journal of Economics*, 15, 69–84.

Cox, C., A. Karam, and R. Murphy (2017): "Social preferences and cooperation in simple social dilemma games," *Journal of Behavioral and Experimental Economics*, 69, 1–3.

Cox, C. A. (2017): "Rent-Seeking and Competitive Preferences," *Journal of Economic Psychology*, 63, 102–116.

Das, R., J. E. Hanson, J. O. Kephart, and G. Tesauro (2001): "Agent-Human Interactions in the Continuous Double Auction," in *Proceedings of the International Joint Conferences on Artificial Intelligence (IJCAI)*.

Davis, A. M., E. Katok, and A. M. Kwasnica (2011): "Do Auctioneers Pick Optimal Reserve Prices?" *Management Science*, 57, 177–192.

Davis, D. D., G. W. Harrison, and A. W. Williams (1993): "Convergence to Nonstationary Equilibria: An Experimental Analysis," *Journal of Economic Behavior and Organization*, 22, 305–326.

Davis, D. D. and A. W. Williams (1986): "The Effect of Rent Asymmetries in Posted Offer Markets," *Journal of Economic Behavior and Organization*, 7, 303–316.

——— (1991): "The Hayek Hypothesis in Experimental Auctions: Institutional Effect and Market Power," *Economic Inquiry*, 29, 261–274.

De Luca, M. and D. Cliff (2011): "Human-Agent Auction Interactions: Adaptive-Aggressive Agents Dominate," in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI)*.

de Melo, C. M. and J. Gratch (2015): "People Show Envy, Not Guilt, when Making Decisions with Machines," in *Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*.

Di Mauro, C. and M. F. Castro (2011): "Kindness, confusion, or ... ambiguity?" *Experimental Economics*, 14, 611–633.

Dua, N., J. S. Heywood, and G. Ye (2013): "Strategic delegation in an experimental mixed duopoly," *Journal of Economic Behavior and Organization*, 87, 91–100.

Duersch, P., A. Kolb, J. Oechssler, and B. C. Schipper (2010): "Rage against the machines: how subjects play against learning algorithms," *Economic Theory*, 43, 407–430.

Duffy, J. (2001): "Learning to speculate: Experiments with artificial and real agents," *Journal of Economic Dynamics & Control*, 25, 295–319.

Duffy, J. and H. Xie (2016): "Group size and cooperation among strangers," *Journal of Economic Behavior and Organization*, 126, 55–74.

Eckel, C. C. and R. K. Wilson (2007): "Social learning in coordination games: Does status matter?" *Experimental Economics*, 10, 317–329.

Embrey, M., G. R. Frèchette, and S. F. Lehrer (2015): "Bargaining and Reputation: An Experiment on Bargaining in the Presence of Behavioural Types," *Review of Economic Studies*, 82, 608–631.

European Commission (2019): *Ethics Guidelines for Trustworthy AI*, High-Level Expert Group on Artificial Intelligence, set up by the European Commission.

Eyster, E. (2019): "Errors in Strategic Reasoning," in *Handbook of Behavioral Economics – Foundations and Applications*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, North-Holland, Amsterdam, vol. 2, chap. 3, 187–259.

Farjam, M. and O. Kirchkamp (2018): "Bubbles in hybrid markets – How expectations about algorithmic trading affect human trading," *Journal of Economic Behavior and Organization*, 146, 248–269.

Fehr, E. and J.-R. Tyran (2001): "Does Money Illusion Matter?" *American Economic Review*, 91, 1239–1262.

——— (2007): "Money illusion and coordination failure," *Games and Economic Behavior*, 58, 246–268.

Ferraro, P. J., D. Rondeau, and G. L. Poe (2003): "Detecting Other-Regarding Behavior with Virtual Players," *Journal of Economic Behavior and Organization*, 51, 99–109.

Ferraro, P. J. and C. A. Vossler (2010): "The Source and Significance of Confusion in Public Goods Experiments," *The B.E. Journal of Economic Analysis & Policy*, 10.

Fox, J. (1972): "The Learning of Strategies in a Simple, Two-Person Zero-Sum Game without Saddlepoint," *Behavioral Science*, 17, 300–308.

Fudenberg, D. and D. K. Levine (1998): *The Theory of Learning in Games*, MIT Press, Cambridge.

Ghosh, S., A. Heifetz, and R. Verbrugge (2016): "Do players reason by forward induction in dynamic perfect information games?" in *Proceedings of the Fifteenth Conference on Theoretical Aspects of Rationality and Knowledge, TARK 2015*, ed. by R. Ramanujam, vol. 215 of *Electronic Proceedings in Theoretical Computer Science, EPTCS*, 159–175.

Ghosh, S., A. Heifetz, R. Verbrugge, and H. de Weerd (2017): "What Drives People's Choices in Turn-Taking Games, if not Game-Theoretic Rationality?" in *Proceedings of the Sixteenth Conference on Theoretical Aspects of Rationality and Knowledge, TARK 2017*, ed. by J. Lang, vol. 251 of *Electronic Proceedings in Theoretical Computer Science, EPTCS*, 265–284.

Gode, D. K. and S. Sunder (1993): "Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality," *Journal of Political Economy*, 101, 119–137.

Gogoll, J. and M. Uhl (2016): "Automation and Morals – Eliciting Folk Intuitions," mimeo, Technical University of Munich.

Gong, M., J. Baron, and H. Kunreuther (2009): "Group cooperation under uncertainty." *Journal of Risk and Uncertainty*, 39, 251–270.

——— (2013): "Why do groups cooperate more than individuals to reduce risks?" *Theory and Decision*, 75, 101–116.

Grebe, T., J. Schmid, and A. Stiehler (2008): "Do individuals recognize cascade behavior of others? – An experimental study," *Journal of Economic Psychology*, 29, 197–209.

Griessinger, T., G. Coricelli, and M. Khamassi (2018): "Sophisticated learning in social competitive interactions," Mimeo.

Grosskopf, B. and R. Sarin (2010): "Is Reputation Good or Bad? An Experiment," *American Economic Review*, 100, 2187–2204.

Guala, F. (2005): *The Methodology of Experimental Economics*, Cambridge University Press, Cambridge.

Guillen, P., E. Fatas, and P. Brãnas Garza (2010): "Inducing efficient conditional cooperation patterns in public goods games, an experimental investigation," *Journal of Economic Psychology*, 31, 872–833.

Hanaki, N., N. Jacquemet, S. Luchini, and A. Zylbersztejn (2016): "Cognitive ability and the effect of strategic uncertainty," *Theory and Decision*, 81, 101–121.

Harrison, G. W. (1989): "Theory and Misbehavior of First-Price Auctions," *American Economic Review*, 79, 749–762.

Harrison, G. W. and K. A. McCabe (1996): "Expectations and Fairness in a Simple Bargaining Experiment," *International Journal of Game Theory*, 25, 303–327.

Harrison, G. W. and M. McKee (1985): "Monopoly Behavior, Decentralized Regulation, and Contestable Markets: An Experimental Evaluation," *The RAND Journal of Economics*, 16, 51–69.

Hendershott, T., C. M. Jones, and A. J. Menkveld (2011): "Does Algorithmic Trading Improve Liquidity?" *The Journal of Finance*, 66, 1–33.

Hertz, N. and E. Wiese (2019): "Good Advice Is Beyond All Price, but What if It Comes From a Machine?" *Journal of Experimental Psychology: Applied*, 25, 386–395.

Houser, D. and R. Kurzban (2002): "Revisiting Kindness and Confusion in Public Goods Experiments," *American Economic Review*, 92, 1062–1069.

Isaac, R. M. and V. L. Smith (1985): "In Search of Predatory Pricing," *Journal of Political Economy*, 93, 320–345.

Ivanov, A., D. Levin, and M. Niederle (2010): "Can Relaxation of Beliefs Rationalize the Winner's Curse?: An Experimental Study," *Econometrica*, 78, 1435–1452.

Iyengar, R. and A. Schotter (2008): "Learning under supervision: an experimental study," *Experimental Economics*, 11, 154–173.

Jaworski, T. and E. O. Kimbrough (2016): "Bubbles, Crashes, and Endogenous Uncertainty in Linked Asset and Product Markets," *International Economic Review*, 57, 155–176.

Johnson, E. J., C. F. Camerer, and T. Rymon (2002): "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining," *Journal of Economic Theory*, 104, 16–47.

Kagel, J. H. (1995): "Auctions: A Survey of Experimental Research," in *The Handbook of Experimental Economics*, ed. by J. H. Kagel and A. E. Roth, Princeton University Press, Princeton, vol. 1, 501–585.

Kagel, J. H. and D. Levin (2001): "Behavior in Multi-Unit Demand Auctions: Experiments with Uniform Price and Dynamic Vickrey Auction," *Econometrica*, 69, 413–454.

——— (2017): "Auctions," in *The Handbook of Experimental Economics*, ed. by J. H. Kagel and A. E. Roth, Princeton University Press, Princeton, vol. 2, 563–637.

Kalayci, K. (2015): "Price complexity and buyer confusion in markets," *Journal of Economic Behavior and Organization*, 111, 154–168.

Kalayci, K. and J. Potter (2011): "Buyer confusion and market prices," *International Journal of Industrial Organization*, 29, 14–22.

Kirchkamp, O. and C. Strobel (2019): "Sharing Responsibility with a Machine," *Journal of Behavioral and Experimental Economics*, 80, 25–33.

Kiyotaki, N. and R. Wright (1989): "On Money as a Medium of Exchange," *Journal of Political Economy*, 97, 927–954.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982): "Rational cooperation in the finitely repeated prisoners' dilemma," *Journal of Economic Theory*, 27, 245–252.

Lapuschkin, S., S. Wäldchen, A. Binder, G. Montavon, W. Samek, and K.-R. Müller (2019): "Unmasking Clever Hans predictors and assessing what machines really learn," *Nature Communications*, 10, 1096.

Levitt, S. D., J. A. List, and D. H. Reiley (2010): "What Happens in the Field Stays in the Field: Exploring Whether Professionals Play Minimax in Laboratory Experiments," *Econometrica*, 78, 1413–1434.

Mago, S. D. (2010): "Costly Buyer Search in a Differentiated Products Model: AN Experimental Study," *Southern Economic Journal*, 77, 282–312.

Manistersky, E., R. Lin, and S. Kraus (2014): "The Development of the Strategic Behavior of Peer Designed Agents," in *Language, Culture, Computation. Computing - Theory and Technology*, ed. by N. Dershowitz and E. Nissan, Springer, Berlin and Heidelberg, 180–196.

March, C. and A. Ziegelmeyer (2018): "Excessive Herding in the Laboratory: The Role of Intuitive Judgments," Cesifo working paper series no. 6855, CESifo, Munich.

McCabe, K., D. Houser, L. Ryan, V. Smith, and T. Trouard (2001): "A functional imaging study of cooperation in two-person reciprocal exchange," *Proceedings of the National Academy of Sciences*, 98, 11832–5.

Meijering, B., H. van Rijn, N. A. Taatgen, and R. Verbrugge (2012): "What Eye Movements Can Tell about Theory of Mind in a Strategic Game," *PloS ONE*, 7, e45961.

Merlo, A. and A. Schotter (1999): "A Surprise-Quiz View of Learning in Economic Experiments," *Games and Economic Behavior*, 28, 25–54.

——— (2003): "Learning by not doing: an experimental investigation of observational learning," *Games and Economic Behavior*, 42, 116–136.

Messick, D. M. (1967): "Interdependent Decision Strategies in Zero-Sum Games: A Computer-Controlled Study," *Behavioral Science*, 12, 33–48.

Moisan, F. and C. Gonzales (2017): "Security under Uncertainty: Adaptive Attackers Are More Challenging to Human Defenders than Random Attackers," *Frontiers in Psychology*, 8.

Morgan, J., H. Orzen, and M. Sefton (2006): "An experimental study of price dispersion," *Games and Economic Behavior*, 54, 134–158.

Neral, J. and J. Ochs (1992): "The Sequential Equilibrium Theory of Reputation Building: A Further Test," *Econometrica*, 60, 1151–1169.

OECD (2019a): *Artificial Intelligence in Society*, OECD Publishing, Paris.

——— (2019b): *OECD Employment Outlook 2019: The Future of Work*, OECD Publishing, Paris.

Oskamp, S. (1971): "Effects of programmed strategies on cooperation in the Prisoner's Dilemma and other mixed-motive games," *Journal of Conflict Resolution*, 15, 225–259.

Petersen, L. and A. Winn (2014): "Does Money Illusion Matter?: Comment," *American Economic Review*, 104, 1047–1062.

Raj, S. B. E. and P. A. Annie (2011): "Analysis on credit card fraud detection methods," in *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*.

Roth, A. E. and J. K. Murnighan (1978): "Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma," *Journal of Mathematical Psychology*, 17, 189–198.

Roth, A. E. and F. Schoumaker (1983): "Expectations and Reputation in Bargaining: An Experimental Study," *American Economic Review*, 73, 362–372.

Schotter, A. (2003): "Decision Making with Naive Advice," *American Economic Review*, 93, 196–201.

Searle, J. R. (1980): "Minds, brains, and programs," *Behavioral and Brain Sciences*, 3, 417–457.

Shachat, J. and J. T. Swarthout (2004): "Do we detect and exploit mixed strategy play by opponents?" *Mathematical Methods of Operations Research*, 59, 359–373.

——— (2012): "Learning about learning in games through experimental control of strategic interdependence," *Journal of Economic Dynamics & Control*, 36, 383–402.

Shapiro, D. A. (2009): "The role of utility interdependence in public good experiments," *International Journal of Game Theory*, 38, 81–106.

Shubik, M., G. Wolf, and B. Poon (1974): "Perception of Payoff Structure and Opponent's Behavior in Related Matrix Games," *Journal of Conflict Resolution*, 18, 646–655.

Silver, D., J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, H. Fan, L. Sifre, G. v. d. Driessche, T. Graepel, and D. Hassabis (2017): "Mastering the game of Go without human knowledge," *Nature*, 550, 354–359.

Smith, V. L. (1962): "An Experimental Study of Competitive Market Behavior," *Journal of Political Economy*, 70, 111–137.

Spiliopoulos, L. (2013): "Strategic adaptation of humans playing computer algorithms in a repeated constant-sum game," *Autonomous Agents and Multi-Agent Systems*, 27, 131–160.

Tesauro, G. and R. Das (2001): "High-performance bidding agents for the continuous double auction," in *Proceedings of the Third ACM Conferences on e-Commerce*, 206–209.

Teubner, T., M. Adam, and R. Riordan (2015): "The Impact of Computerized Agents on Immediate Emotions, Overall Arousal and Bidding Behavior in Electronic Auctions," *Journal of the Association of Information Systems*, 16, 838–879.

Topol, E. J. (2019): "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine*, 25, 44–56.

Upadhyay, A. K. and K. Khandelwal (2018): "Applying artificial intelligence: implications for recruitment," *Strategic HR Review*, 17, 255–258.

van den Bos, W., J. Li, T. Lau, E. Maskin, J. D. Cohen, P. R. Montague, and S. M. McClure (2008): "The value of victory: social origins of the winner's curse in common value auctions," *Judgment and Decision Making*, 3, 483–492.

Veiga, H. and M. Vorsatz (2009): "Price manipulation in an experimental asset market," *European Economic Review*, 53, 327–342.

——— (2010): "Information aggregation in experimental asset markets in the presence of a manipulator," *Experimental Economics*, 13, 379–398.

Walker, J. M., V. L. Smith, and J. C. Cox (1987): "Bidding Behavior in First Price Sealed Bid Auctions: Use of Computerized Nash Competitors," *Economics Letters*, 23, 239–244.

West, R. L. and C. Lebiere (2001): "Simple games as dynamic, coupled systems: randomness and other emergent properties," *Journal of Cognitive Systems Research*, 1, 221–239.

Winter, E. and S. Zamir (2005): "An Experiment with Ultimatum Bargaining in a Changing Environment," *Japanese Economic Review*, 56, 363–385.

Wolf, G. and M. Shubik (1978): "Market Structure, Opponent Behavior, and Information in a Market Game," *Decision Sciences*, 9, 421–428.

Wu, C.-T., Y.-T. Fan, Y.-R. Du, T.-T. Yang, H.-L. Liu, N.-S. Yen, S.-H. Chen, and R.-M. Hsung (2018): "How Do Acquired Political Identities Influence Our Neural Processing toward Others within the Context of a Trust Game?" *Frontiers in Human Neuroscience*, 12.

Yamakawa, T. and T. Okano, Yoshitaka ans Saijo (2016): "Detecting motives for cooperation in public goods experiments," *Experimental Economics*, 19, 500–512.

BERG Working Paper Series (most recent publications)

121 Marco **Sahm**, Are Sequential Round-Robin Tournaments Discriminatory?, March 2017

122 Noemi **Schmitt**, Jan **Tuinstra** and Frank **Westerhoff**, Stability and welfare effects of profit taxes within an evolutionary market interaction model, May 2017

123 Johanna Sophie **Quis** and Simon **Reif**, Health Effects of Instruction Intensity – Evidence from a Natural Experiment in German High-Schools, May 2017

124 Lisa **Planer-Friedrich** and Marco **Sahm**, Strategic Corporate Social Responsibility, May 2017

125 Peter **Flaschel**, Matthieu **Charpe**, Giorgos **Galanis**, Christian R. **Proaño** and Roberto **Veneziani**, Macroeconomic and Stock Market Interactions with Endogenous Aggregate Sentiment Dynamics, May 2017

126 Christian **Menden** and Christian R. **Proaño**, Dissecting the Financial Cycle with Dynamic Factor Models, May 2017

127 Christoph **March** and Marco **Sahm**, Contests as Selection Mechanisms: The Impact of Risk Aversion, July 2017

128 Ivonne **Blaurock**, Noemi **Schmitt** and Frank **Westerhoff**, Market entry waves and volatility outbursts in stock markets, August 2017

129 Christoph **Laica**, Arne **Lauber** and Marco **Sahm**, Sequential Round-Robin Tournaments with Multiple Prizes, September 2017

130 Joep **Lustenhouwer** and Kostas **Mavromatis**, Fiscal Consolidations and Finite Planning Horizons, December 2017

131 Cars **Hommes** and Joep **Lustenhouwer**, Managing Unanchored, Heterogeneous Expectations and Liquidity Traps, December 2017

132 Cars **Hommes**, Joep **Lustenhouwer** and Kostas **Mavromatis**, Fiscal Consolidations and Heterogeneous Expectations, December 2017

133 Roberto **Dieci**, Noemi **Schmitt** and Frank **Westerhoff**, Interactions between stock, bond and housing markets, January 2018

134 Noemi **Schmitt**, Heterogeneous expectations and asset price dynamics, January 2018

135 Carolin **Martin** and Frank **Westerhoff**, Regulating speculative housing markets via public housing construction programs: Insights from a heterogeneous agent model, May 2018

136 Roberto **Dieci**, Noemi **Schmitt** and Frank **Westerhoff**, Steady states, stability and bifurcations in multi-asset market models, July 2018

137 Steffen **Ahrens**, Joep **Lustenhouwer** and Michele **Tettamanzi**, The Stabilizing Role of Forward Guidance: A Macro Experiment, September 2018

138 Joep **Lustenhouwer**, Fiscal Stimulus in an Expectation Driven Liquidity Trap, September 2018

139 Tim **Hagenhoff**, An aggregate welfare optimizing interest rate rule under heterogeneous expectations, October 2018

140 Johanna Sophie **Quis**, Anika **Bela** and Guido **Heineck**, Preschoolers' self-regulation, skill differentials, and early educational outcomes, December 2018

141 Tomasz **Makarewicz,** Traders, forecasters and financial instability: A model of individual learning of anchor-and-adjustment heuristics, January 2019

142 Carolin **Martin**, Noemi **Schmitt** and Frank **Westerhoff**, Housing markets, expectation formation and interest rates, January 2019

143 Marc P. **Saur**, Markus G. **Schlatterer** and Stefanie Y. **Schmitt**, Horizontal product differentiation with limited attentive consumers, January 2019

144 Tim **Hagenhoff** and Joep **Lustenhouwer**, The Rationality Bias, February 2019

145 Philipp **Mundt** and Ilfan **Oh**, Asymmetric competition, risk, and return distribution, February 2019

146 Ilfan **Oh**, Autonomy of Profit Rate Distribution and Its Dynamics from Firm Size Measures: A Statistical Equilibrium Approach, February 2019

147 Philipp **Mundt**, Simone **Alfarano** and Mishael **Milakovic**, Exploiting ergodicity in forecasts of corporate profitability, March 2019

148 Christian R. **Proaño** and Benjamin **Lojak**, Animal Spirits, Risk Premia and Monetary Policy at the Zero Lower Bound, March 2019

149 Christian R. **Proaño**, Juan Carlos **Peña** and Thomas **Saalfeld**, Inequality, Macroeconomic Performance and Political Polarization: An Empirical Analysis, March 2019

150 Maria Daniela **Araujo P.**, Measuring the Effect of Competitive Teacher Recruitment on Student Achievement: Evidence from Ecuador, April 2019

151 Noemi **Schmitt** and Frank **Westerhoff**, Trend followers, contrarians and fundamentalists: explaining the dynamics of financial markets, May 2019

152 Yoshiyuki **Arata** and Philipp **Mundt**, Topology and formation of production input interlinkages: evidence from Japanese microdata, June 2019

153 Benjamin **Lojak**, Tomasz **Makarewicz** and Christian R. **Proaño**, Low Interest Rates, Bank's Search-for-Yield Behavior and Financial Portfolio Management, October 2019

154 Christoph **March**, The Behavioral Economics of Artificial Intelligence: Lessons from Experiments with Computer Players, November 2019