PSYCHOLOGICAL GAME THEORY: A REVIEW OF CURRENT LITERATURE*

DANIEL NEICU**

Abstract

Experimental evidence appears to contradict traditional game theory predictions in numerous settings. Although a solid basis for characterizing equilibria, game theory needed to – and to some extent did – go through some major developments and transgress the barriers between different social sciences to more accurately depict human behaviour in economic circumstances. Psychological game theory, first developed in the 80s, is one such advancement. By integrating beliefs directly within utility, it seeks to tie affective psychology to economic choice and thus better predict decision making. This paper reviews existing literature on psychological game theory and a strand of prospect theory relevant to emotional decision-making, and proposes further refinements and possible extensions.

JEL codes: D03, D84.

Keywords: belief-dependent decision making; dynamic psychological games; emotions; prospect theory; psychological game theory

I. INTRODUCTION

Over the past decades, the strict, law-like constructs of neoclassicism in economics have come under pressure due to experimental evidence of *unpredicted* behaviour. Behavioural economists have tried to fill this gap in some areas by adjusting neoclassical theories via a formalisation of experimental results. However, as many have pointed out, behavioural

* I am grateful to Christophe Crombez, Georg Kirchsteiger, Martin Dufwenberg, Pierpaolo Battigalli, Martin Kaae Jensen and Michalis Drouvelis for their useful comments on various versions of this paper.

** Daniel Neicu, Address: Department of Managerial Economics, Strategy and Innovation, KU Leuven, Naamsestraat 69, 3000 Leuven, Tel. +32 16 326700, E-mail: daniel.neicu@econ.kuleuven.be.

economics is not trying to stamp out neoclassical theory, but rather increase its realism and prediction power by incorporating psychological underpinnings.

Although some regard it as a new field, behavioural economics relates to work done in the 50s through to the 70s by social scientists like Simon (1955), Becker (1968), or Tversky & Kahneman (1974, 1979), some of it as a critique to neoclassical models (Allais 1953). The theory of consumer choice proposed by Thaler (1980) started a new movement which quickly became a standard in marketing. Nowadays there is an equally important amount of research focusing on the production and supply side of economics, where companies have previously been thought to act with unmitigated rationality due to the competitive environment they operate in, with mistakes or deviations from rationality being fatal in the long run.

Arguably at the basis of current neoclassical concepts lies game theory, which has been developed in order to portend how different agents will act in strategic encounters. Economic decision making is a prime target for such predictions, partially due to the rigorous solution concepts offered by the framework.

However, some of the predictions of traditional game theory have been contradicted by experimental results on numerous occasions. Forsythe et al. (1994) and Camerer (2003) provide an overview of experimental evidence for situations in which traditional equilibria predictions are not met. Although not extensive, here are some of their findings.

In ultimatum games, experiments have shown that in around half of the cases the proposer offers on average 40% of the entire sum to the responder, while responders reject offers smaller than 20% of the total, whereas the equilibrium predicted by traditional game theory is that the proposer offers nothing and the responder accepts it.[1, 2]

The results in ultimatum games seem to be consistent to some extent by experiments with dictator games, in which the proposer has unmitigated decision power.[3] Some dictator experiments with fixed splits have shown that in almost 75% of cases people split the sum 50–50. However, when the players are left to choose themselves the split, the average offer is still 20%, well above the predicted 0% in traditional game theory.

In prisoner's dilemma experiments players cooperated 50% of the time in one-shot games. However, if the experiments are done with random choice of partners, people tend to cooperate less, but there remains a set of players who almost always cooperate, no matter who their partner is. For reference, traditional theory predicts that both players defect, as cooperation is not a best reply to defection.

1 Ultimatum games are simple two-player frameworks in which one player offers to split a sum of money, and the other player can either accept or reject the proposal, after which the sum is distributed as agreed and the game ends. If player 2 rejects the proposal, both players get nothing.

2 Median offers are around 40–50% and average offers around 30–40%.

3 Dictator games are similar to ultimatum games, but only one player can make an "offer", which is implemented and the game ends without the second player making any decision.

Furthermore, in public goods games results show that on average, players contribute half of their endowment towards public goods. If the same randomization is applied as above, contributors become less numerous and punishment and free-riding effects appear.

In experiments where the sums of money involved were raised in order to check for methodological bias, researchers have found that raising the stake of a game by a large magnitude only slightly changes the results, which thus remain robust to methodological bias.

These examples show that people might not be strict maximizers of "egoistic" utility. Instead, they exhibit social preferences towards fairness or reciprocity, or experience emotions such as guilt or envy. Hence, the results can be explained by allowing utility to reflect social preferences, at the same time keeping the rationality assumption that people maximise different forms of utility.

These recurrent results have kick-started several behavioural models that make use of psychological insights in order to better anticipate outcomes under specific conditions.

The objective of this paper is to present a comprehensive review of the current literature on one of these models, namely psychological game theory, as well as proposing its further development by including prospect theory concepts in order to improve predictive power.

The paper is structured as follows. Section 2 presents an overview of past and current relevant research, specifically focusing on psychological games (section 2.1), dynamic psychological games (section 2.2) and reference-dependent preferences (section 2.3). Building on this, in section 3 I propose some paths for further research, and conclude with a brief subsection discussing the possible results of a new model. Section 4 concludes the paper.

II. CURRENT STATE OF RESEARCH

The experiments mentioned above are at the basis of new decision making models, one set of which is known as psychological game theory. Psychological game theory captures an array of conducts by means of a belief system in which the assumption of rationality is relaxed, but not waived, as players' emotional responses can be modelled by incorporating beliefs into utility functions. Psychological evidence in support of belief-dependent emotions comes from Elster (1998), who describes a series of emotions and argues that they are triggered by people's beliefs. The same issue is tackled by Loewenstein et al. (2001) in a paper where belief-related feelings are analysed in a two-way structure of anticipated vs. anticipatory emotions, both dependent on the subjects' beliefs.

A. PSYCHOLOGICAL GAMES

Geanakoplos, Pearce & Stacchetti (1989) (hereafter GPS) were among the first to develop a model of belief-dependent strategies.[4] Their framework defines utility as being directly conditioned by players' actions and by their initial beliefs about how the game will be played. The reason for this is that, intuitively, people make decisions based not only on the monetary outcome of their actions, but on a more complex form of utility which might be influenced by, among others, beliefs. Prior expectations of an outcome and the failure to meet them can trigger a wide range of emotional responses, leading to what in common language is referred to as irrational behaviour. However, GPS show that rationality need not be forgone in this context, but rather reinforced in order to explain seemingly irrational play. Moreover, due to the dependency of (some) emotions on beliefs, psychological utilities that players get at terminal nodes of a game are endogenous and need to be defined accordingly.

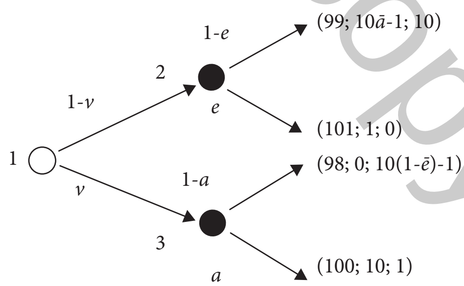
The stated purpose of GPS is to develop a comprehensive framework that can be applied to a wide variety of games in the sense that it will allow more accurate predictions of game play by means of redefined equilibria. An important characteristic of the framework is that it allows the use of existing equilibrium concepts, although backward induction, among others, cannot be applied in psychological games.

The disjunction with traditional game theory lies in the fact that rationality assumptions incorporate beliefs directly into utility functions, which are defined as

$$u_i : \Sigma \times \bar{M}_i \rightarrow \mathbb{R}$$ [5]

In order to illustrate the need of a more complex domain of utility that includes beliefs, GPS use the following 3-player game.

Figure 1. Psychological game reproduced from GPS (1989: p. 73)



[4] GPS acknowledge previous work by Gilboa & Schmeidler (1988) on information-dependent games, where utilities depend on ex ante information about the approximate set of possible outcomes. These sets endogenise utilities in a similar fashion as hierarchies of beliefs do in GPS.

[5] Where u_i is player i's utility, Σ is the set of strategy profiles, and $\bar{M}_i$ is player i's hierarchy of initial beliefs about the others' strategies and initial beliefs (first and higher-order).

The payoffs of players 2 and 3 include terms $\bar{a}$ and $\bar{e}$, which are, respectively, player 2's expectation of the probability that player 3 would play down if given the move, and player 3's expectation of the probability that player 2 would play down if his node were to be reached. In other words, player 2 and 3's utilities comprise the satisfaction that each would get from hurting player 1 (by diminishing his payoff), which is proportional to the (monetary) payoff lost by each due to 1's observed action. An example of game play is the following: if player 1 chooses down, and player 3 believes that if player 1 had chosen up then player 2 would have chosen up ($\bar{e} = 0$) and she (player 3) would have gotten payoff 10, then she will choose up and hurt player 1. Had she thought player 2 would have played down ($\bar{e} = 1$), she would choose down and reward player 1.[6] This psychological game structure shows that backward induction is made impossible due to players' intentions at end nodes being dependent on their beliefs $\bar{a}$ and $\bar{e}$. There are two subgame perfect psychological equilibria, which will be defined later, in mixed strategies, when $\bar{e} = e = 4/5$, $\bar{a} = a = 1/5$ and $v = 0$.

In order to define equilibria, the following notations are employed:

- σ_i is one of player i's behaviour strategies while σ_{-i} is a profile of strategies of all players but i;
- $\sigma = (\sigma_1, \sigma_2, \ldots \sigma_n)$ is a strategy profile;
- $\bar{M}$ is a profile of initial belief hierarchies (of all players).

GPS define a psychological Nash equilibrium a pair $\left(\bar{M},\hat{\sigma}\right)$ of a normal form game with the following properties:

i. $u_i\left(\bar{M},\hat{\sigma}\right) \geq u_i\left(\bar{M}, \sigma_i,\widehat{\sigma_{-i}}\right)$, for all i and all σ_i, and

ii. $\bar{M}$ does not contradict strategy profile $\hat{\sigma}$, which is to say that every player believes with probability 1 that every other player plays $\widehat{\sigma_{-i}}$, and that every other player j believes that every other player plays $\widehat{\sigma_{-j}}$ and so forth.

Moreover, $\left(\bar{M},\hat{\sigma}\right)$ is a subgame perfect psychological equilibrium if it is a psychological Nash equilibrium and $\hat{\sigma}$ is a subgame perfect equilibrium, given $\bar{M}$.

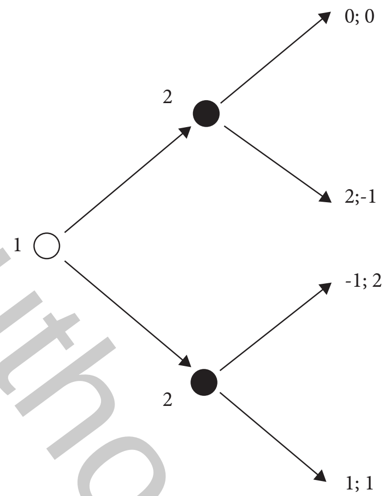
It follows that if a psychological game has continuous payoff functions, it has a sub-game perfect psychological equilibrium.

Although it allows for belief-dependent motivation, this model still relies heavily on the rationality assumption, each player being rational and knowing that the others are also rational. GPS admit that an important drawback is that beliefs are static, whereas intuitively they should be updated after each action has taken place. Although they use a type of updated beliefs in sequential equilibria, they lack the formal constructs for capturing updated beliefs about others' beliefs. Furthermore, other players' ($-i$) beliefs do not enter the domain of player i's utility, while strategies influence utilities only by their impact on end-nodes, although there might be reasons to believe they should also influence intermediate nodes.

[6] Note that the fact that player 3 chooses her highest payoff does not imply that she does not have other-regarding preferences; they are incorporated in her utility function by means of belief $\bar{e}$.

In order to demonstrate the drawbacks of static belief hierarchies, consider the following extensive form game.

Figure 2. Extensive form game with consistent initial beliefs



In this example, if player 2 has consistent initial beliefs, then strategy combination (*Down*, *down down*) forces player 2 to consider player 1 as altruistic even if he encounters the unexpected move *Up*. Because payoffs depend only on consistent initial beliefs that do not contradict the strategy combination, off the equilibrium path beliefs remain unchanged (at their initial state), and thus the need for dynamic belief updating.[7]

B. DYNAMIC PSYCHOLOGICAL GAMES (DPGS)

One of the few extensions of this model comes from Battigalli & Dufwenberg (2009), who address some issues of the original framework.[8] The next sub-sections are a description of their method.

1. Updated conditional beliefs

Firstly, their model allows for updated beliefs, where players "revise" them at each node based on the actions that have been taken by every (other) player in order for that node to be

[7] On the equilibrium path beliefs do not change either.

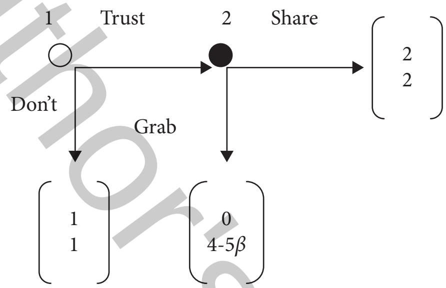
[8] Battigalli & Dufwenberg (2009) argue that only a handful of articles have attempted to extend the overall framework of GPS, such as Kolpin (1992), or Segal & Sobel (2007).

reached. The intuitive argument for this is illustrated by a simple game with two players, with the following extensive form.

The game is used to model player 2's sentiment of guilt towards player 1. The intuition is that player 2 can *let player* 1 *down* if player 1's material payoff will be less than what she expected. Player 1 assigns a probability α to player 2's strategy *Share if Trust*. As this probability is not known by player 2, he has a belief (or expectation) β about it.[9] This belief is conditional on player 1 choosing *Trust*.[10] If player 2 does not want to feel guilty by letting down player 1, then his utility will depend directly on β, as depicted in the figure. Note that β is an updated belief about α, and is conditional on player 1 choosing *Trust*.[11]

To illustrate prediction in an updated-beliefs system, Battigalli & Dufwenberg (2009) construct the following reasoning, depicted as *psychological forward induction*:

Figure 3. Psychological game reproduced from Battigalli & Dufwenberg (2009, p. 6)



assuming that player 1 chooses *Trust*, she must have done so because she believes that player 2 will choose *Share* with probability $\alpha \geq 0.5$.[12] Player 2, being rational, anticipates this, so that $\beta \geq 0.5$, and thus $(4–5\beta) < 2$, which compels him to choose *Share* for the bigger utility (2). As player 1 is also rational, she can figure this out, so she chooses *Trust*, expecting player 2 to choose *Share* ($\alpha = 1$). This leads the rational player 2 to believe that player 1 believes he will choose *Share* (now $\beta = 1$), so his utility from sharing becomes even greater than the one from grabbing (2 > -1). All these deductions lead to the prediction of the path (*Trust*, *Share*).

In order to summarize this framework, some of its notation, used to describe an extensive form game, needs to be introduced:

- $N = \{1, \ldots, n\}$ is the player set;
- a history $h = (a^1, \ldots, a^k)$ is a sequence of action profiles a^t describing the actions chosen by each player $i \in N$ at stage t;
- H is a finite set of feasible histories h;

9 Parameters α and β are point beliefs.

10 The meaning of conditional beliefs is explained below.

11 Terminal nodes describe utilities, not material payoffs for this game.

12 Based on Dufwenberg (2002).

- $A_i(h)$ represents the set of feasible actions of player i at history h;
- Z is the set of payoffs at end-nodes;
- $\zeta(s) \in Z$ is the material payoff (terminal history) induced by strategy profile s;
- S_i is the set of player i's pure strategies.

Updated beliefs are an essential extension to GPS, as they take the framework a step further towards depicting decision-making behaviour under the assumption of belief-dependent motivation. Although highly intuitive, describing a system of updated beliefs is cumbersome and requires complex notations. Battigalli & Dufwenberg (2009) build their model following Battigalli & Siniscalchi's (1999) theory of hierarchies of conditional beliefs. They define a conditional probability system as a function of the form $\mu(\cdot|\cdot): \mathcal{B} \times C \rightarrow [0,1]$, where for all $E \in \mathcal{B}\, and\, F, F' \in C$ we have

i. $\mu(\cdot|F) \in \Delta(X)$,

ii. $\mu(F|F) = 1$,

iii. $E \subseteq F' \subseteq F$ implies $\mu(E|F) = \mu(E|F')\mu(F'|F)$,

where the set X is a compact Polish space of strategies, $\Delta(X)$ is the set of all probabilities of events in the Borel sigma-algebra $\mathcal{B}$ of X, $C \subseteq \mathcal{B}$ is the set of potentially observable events (conditioning events corresponding to histories h), and E, F and F' are events in $\mathcal{B}$ and C, respectively.

In other words, C is a collection of observable events in the set X of possible events (or actions, or histories). Furthermore, for each observed event $F \in C$, player i attributes a probability, and all such probability measures make up $\Delta(X)$. The probability of an event F conditional on its own observation is 1.

It follows that hierarchies of conditional probability systems (cps) are defined recursively as sequences of infinite-order cps, and each k-order cps is a k-order joint belief on the other players' strategies (s_{-i}) and (k-1)-order beliefs.

Finally, a cps μ_i is *coherent* if beliefs of distinct orders assign the same conditional probabilities to lower-order events, and it is collectively coherent if it satisfies coherency in beliefs of any order $k \in \mathbb{Z}_+$.

The above theory on conditional probability systems leads to the definition of M_i, the set of player i's collectively coherent hierarchies of up-to-infinite orders, which Battigalli & Dufwenberg (2009) include in the domain of players' utility functions.

In other words, M is such that at any information set all players believe that all other players believe etc. that everybody has chosen the exact actions leading to that information set with probability 1.[13] In parts of the game which do not lead to that information set, beliefs remain as specified by $\bar{M}$, the players' initial beliefs.

13 Where $M = \prod_{i \in N} M_i$.

2. *Others' beliefs*

Furthermore, Battigalli & Dufwenberg (2009) define a player's utility such that it incorporates not only her own beliefs, but also others' beliefs.[14] Their justification for this is twofold: firstly, it may be an adequate description of a number of social reward systems; secondly, it is easier to model games with others' beliefs integrated in the utility function, as updated second-order beliefs can be modelled as others' first-order beliefs, thus "skipping" one belief order. However, when players are uncertain – and there are intuitive reasons to presume this – about others' beliefs, they have to use probability assessments in order to weigh the different possibilities, which in turn lead to arguably the same complexity as a higher-order belief.

3. *Dependence on plans*

The third extension is waiving the assumption that strategies influence utility only by their impact on terminal nodes. Battigalli & Dufwenberg (2009) argue that many forms of belief-dependent motivation are not captured by terminal nodes, but rather by actions within the game and overall strategies. In other words, a player's expectations about her (material) payoff given her strategy may depend not only on her beliefs about other players' strategies and beliefs, but also on her own game plan. A key factor in this assumption is that motivation may be dependent on *intentions*. As intentions depend both on beliefs and plans, the domain of utility includes conditional beliefs and strategies of other players.[15]

Given the above, the authors consider a wider domain for utilities than traditional games or GPS's framework, using the following (measurable and bounded) *psychological payoff function*, which incorporates terminal nodes (histories), updated conditional beliefs of up-to-infinite order of player i, other players' ($-i$) updated conditional beliefs of up-to-infinite order and their (pure) strategies:

$$u_i : Z \times M_i \times \prod_{j \neq i} (M_j \times S_j) \rightarrow \mathbb{R}.^{16}$$

Subsequently, the definition of a *psychological game* based on the extensive form $\langle N, H \rangle$ is a structure $\Gamma = \langle N, H, u_{i \in N} \rangle$.

14 The intuition that one's utility is dependent on others' beliefs has previously been hypothesized in models of incomplete information by Bernheim (1994), Dufwenberg & Lundholm (2001), Caplin & Leahy (2004), and Kószegi (2006).

15 A plan defines how a player intends to play at every decision node. The difference between a plan and a strategy is that plans are dynamic and updated as play goes on. Plans are indirectly included in utility via beliefs and strategies.

16 They incorporate strategies within the utility function based on the assumption that strategies coincide with plans, i.e. players' beliefs about their behaviour at each node.

4. *Sequential equilibrium*

The equilibrium concept used by Battigalli & Dufwenberg (2009) is based on the notion of sequential equilibrium in traditional game theory due to Kreps & Wilson (1982). It implies that the existence of equilibria in extensive form games requires the definition of players' assessments as profiles of behaviour strategies and conditional first-order beliefs. In DPGs, an assessment is a profile (σ,μ) of behavioural strategy profiles and conditional beliefs.

Following the above, a dynamic psychological game $\Gamma = \langle N, H, u_{i\in N}\rangle$ has a psychological sequential equilibrium (σ^*,μ^*) if at each information set h, $\mu^*(\cdot|h)$ is such that all players believe with probability 1 that strategy profile σ^* is played (consistency) and each player active at h maximizes his expected payoff, given his beliefs $\mu_i(s_{-i}|h)$.

More formally, the assessment (σ,μ) is a sequential equilibrium if it is *consistent* and for all $i \in N, h \in H \setminus Z, s_i^* \in S_i(h)$

$$Pr_{\sigma_i}\left(s_i^*\,|\,h\right) > 0 \Rightarrow s_i^* \in \arg\max_{s_i \in S_i(h)} E_{s_i,\mu_i}[u_i\,|\,h].$$ [17]

Assessment (σ,μ) is *consistent* if first-order beliefs (μ^1) are derived from strategy profile σ, and all higher-order beliefs assign probability 1 to lower-order beliefs.[18]

The second part of the definition states that if there is a positive probability of player i playing s_i^*, conditional on non-terminal history h and given strategy profile σ_i, then strategy s_i^* is one that maximises the expected value of utility u_i, conditional on non-terminal history h, given consistent beliefs in μ_i.

The definition of consistency leads i's strategy profile σ_i to be identical to the first order beliefs of all $j \neq i$ about i. In other words, sequential equilibrium is based on equilibrium in beliefs of first and higher-orders.

Consequently, Battigalli & Dufwenberg (2009) prove that any game $\Gamma = N, H, u_{i\in N}$ has at least one psychological equilibrium if u_i is continuous. Furthermore, if payoffs are constant in M, the psychological sequential equilibrium is equivalent to Kreps & Wilson's (1982) notion of equilibrium.

5. *Non-equilibrium analysis*

A forth essential point in DPGs is that non-equilibrium analysis is just as important as the analysis of equilibrium play. The issue arises due to the strong assumptions that players coordinate on equilibrium strategies, which might not be the case for a large number of games. However, even if players do not coordinate on equilibria, they might do so without the

17 $Pr_{\sigma j}(\cdot|\hat{h})$ is the probability measure over j's strategies conditional on $\hat{h}$ derived from σ_j.

18 Profile μ_j^1 of first-order cps is derived from behavioural strategy profile σ_i if it equals the product of probability measures over players –i's strategies conditional on history h.

loss of rationality, as non-equilibrium coordination can arise in specific situations, even when players are rational and believe that others are rational *ad infinitum*.[19]

The definition of sequential equilibrium assumes that all players hold correct initial and higher-order beliefs, so that *i*'s beliefs about -*i*'s intentions and beliefs are independent of –*i*'s behaviour. This assumption is relaxed in DPGs, where a player who observes an (unexpected) action updates his belief about another player's intention of choosing that action. However, as the opponents' intentions and their perceptions are of utmost importance in DPGs, the framework holds the mechanisms needed to deduce others' beliefs from their observable actions through updated beliefs based on conditional probability systems. It does so by relaxing the consistency-in-beliefs assumption in sequential equilibria, allowing players to update higher-order beliefs when opponents make unexpected choices (out-of-equilibrium moves).

Non-equilibrium analysis relaxes powerful assumptions with minimal loss of predictive power, as notions of forward induction (solution concept described for the game in Figure 1), *rationalizability* and self-confirming equilibrium can predict play. Simple self-confirming equilibrium is defined by players acting sequentially rational and having their beliefs confirmed. Furthermore, if there is initial common belief in sequential rationality and confirmed beliefs, the equilibrium is called *rationalizable self-confirming equilibrium* (Battigalli & Dufwenberg, 2011). These last two forms of equilibria are *work in progress* and come to the aid of prediction under relaxed assumptions. They are leading candidates for further extensions to GPS.

6. *Possible extensions of DPG*

Arguably the DPG framework is a big step towards better understanding and predicting human decision making under the influence of belief-influenced utility. Its system of updated conditional beliefs formalizes intuitions about such behaviour in a way that allows for enough flexibility to provide a solid basis for a wide range of psychological motivations that deviate from standard predictions of game theory.

As discussed in the previous section, equilibrium concepts such as sequential equilibrium might not always lead to a good prediction of play. This is so especially due to strong assumptions as consistency of beliefs. Battigalli & Dufwenberg (2009) take a different path to enhance predictive power of DPGs by analysing non-equilibrium play that may lead to a predicted outcome, which could be seen as a parallel to the theory on *obvious ways* to play games in the traditional framework.

Furthermore, although Battigalli & Dufwenberg (2009) restrict their analysis to multi-stage games with complete information, they do so only for simplifying the notation, as the

[19] Forward induction can sometimes describe obvious ways to play a game without assuming any coordination on equilibria.

framework can be applied to incomplete information games. They also admittedly restrict analysis to rational players, arguing that their general framework could be approached from a learning perspective or by concepts of bounded rationality.

Learning models of decision making are constructed on random choices of strategy at the beginning of play, after which players *learn* to play better strategies based on some learning mechanism. Another distinct feature is that games are played repeatedly, so that learning can occur over time. The following paragraphs describe the most relevant types of learning models from the perspective of DPGs. The discussion of bounded rationality will follow.

Reinforcement learning allows for learning from own experience, such that played strategies get *reinforced* proportionally to their obtained payoff (Erev & Roth 1998). As the model is not analytically tractable, a number of simulations have shown that some games' outcomes can be predicted in the medium or long-run.[20] The main criticisms of this model are that *i)* predictions of the medium and long-run are impossible to test in experiments due to obvious time constraints and *ii)* players do not use counterfactual reasoning. The latter is addressed in experience-weighted attraction models of learning, where random strategies' *attraction* depends both on actual and virtual experience of players.[21] The issues with this second type of models are that it does not allow learning from other players' (observed) experience and that it provides no analytical solution due to a large number of parameters which have to be estimated for simulations.

Markov learning models are distinctively appropriate candidates for extensions of DPG due to their definition of Markov states, which is reminiscent to Battigalli & Dufwenberg's (2009) set of histories H. In these types of models, a state ω_k describes the strategies chosen by every player during the past k rounds. Randomisation of strategies is defined as experimentation (or mistake). The properties of Markov chains applied to these models permit learning (convergence) over a finite number of periods, usually less than reinforcement and experience-weighted attraction learning. Predicting outcome is based on defining stochastically stable ω states, or states which describe strategies that do not disappear (do not reach close to zero probability of being played) in the long run, when experimentation tends to zero. The framework allows for the definition of a number of learning methods, such as imitation learning (Vega-Redondo 1997), best-reply with one-period memory (Huck, Normann & Oechssler 1999), and best-reply with finite memory and random sampling (Young 1993). Imitation learning describes stochastically stable states as those that are played with high probability in the long run due to imitation of other played strategies. Best-reply with one-period memory models involve agents revising their strategy with random non-negative probability at time period t, and choosing a strategy which is a best reply to the other

[20] E.g. experimental results of the ultimatum game show that in round 10 the median offer is 50% of the payoff, while simulations of reinforced learning predict a model offer of 4 after 300 runs and 0 or 1 after 1 million runs (Roth & Erev 1995).

[21] Thus the link with DPGs comes from beliefs about strategies that have not been played, which might provide an interesting – though cumbersome to define – extension of DPGs.

agents' strategies at *t*-1. Best-reply with finite memory and random sampling is a similar model, but one which allows for *m*-round memory out of the past *n* rounds ($m<n$) with random choice (with full support) of *m* over *n*.

A series of experiments have compared the different learning models in order to assess their applicability and test their prediction power in controlled situations. Offerman, Potters & Sonnemans (2002) provide experimental evidence for one-period best reply vs. imitation learning in a simulated market environment consisting of 3 firms, with the game being played over 100 rounds.. By using 3 different information *treatments* regarding demand and cost functions, quantities and prices of different players, they discriminate between the analysed learning models. Their results show that, no matter the information available to players, quantities are close to Cournot levels (but always slightly below). However, the more information players receive, the larger the quantities traded on the virtual market. In this experiment imitation learning predicts Walrasian quantities for all players in equilibrium, whereas one-period best reply predicts the Cournot quantities observed in all information *treatments*. These results follow earlier work by Huck, Normann & Oechssler (1999) in a paper opposing the same types of learning. They discover that when information allows for a wide range of learning models, a large number of players choose some form of best reply, while a minority choose imitation.

The results of these experiments suggest that the type of learning involved in a dynamic process is directly influenced (and induced) by the type of information agents receive during game play. Moreover, they show that learning models can be useful candidates for predicting long-term play on anonymous markets. The fact that agents seem to imitate more if they have more information about their partners' behaviour in previous rounds may imply that there is an intrinsic *imitation* predisposition, which is player specific. Intuitively, the more information about the market one has, the better a player can formulate a best reply strategy. However, this rationality does not correspond to reality (at least experimental reality), leading to the conclusion that the outcome of game play in complex situations is not only contingent on payoff (profit maximisation), and thus utility needs to be redefined, possibly in a DPG type of framework.

Battigalli & Dufwenberg (2009) argue that it might be worthwhile to approach psychological games from a learning point of view. However, they admit that standard intuition about how players learn to play equilibria in repeated games by reaching a point where they hold correct beliefs about the actions of their opponents might not suffice for DPGs. The reason is that actions are induced by hierarchies of beliefs. Therefore, players need to learn not only the others' actions, but also their beliefs. This raises the issue of correct beliefs about beliefs, which is not only mathematically burdensome, but also unsupported from a psychological perspective, as beliefs cannot be assessed objectively by a third party.

Extending the DPG framework by integrating learning models and bounded rationality will undoubtedly introduce the need to redefine solution concepts, as the algebra behind them requires such a change.

7. *Applications of DPGs*

Although other frameworks for modelling emotions such as guilt, reciprocity or envy exist, the use of DPG-like models remains the basis for the analysis of such specific issues.[22] Building on the work of Baumeister, Stillwell & Heatherton (1994), Battigalli & Dufwenberg (2007) assess how guilt plays a role in psychological games. However, they go no further than inferring the presence of guilt from theoretical actions, such as not performing as the others expect. Their model defines utility as a combination of payoffs and initial beliefs about strategies, where the impact of guilt on the actual payoff is measured by how much (monetary) "damage" a player inflicts on another player by playing a given strategy. The measure of damage involves calculating the difference between what a player initially thought he would get as payoff if all the others played a given set of strategies and he plays a strategy s_i, and what he actually gets at the end of play by playing s_i and observing others' strategies. Thus, if she has correct initial beliefs and correct updated beliefs, she will be unharmed by the other players, as she will receive the payoff she expected to get. Further, the paper defines simple guilt – as guilt depending on first order beliefs – and guilt-from-blame – as dependent on second and third order beliefs.

As do most papers in the field, Battigalli & Dufwenberg (2007) assess the utility function that includes guilt by means of a *guilt sensitivity* parameter, of which they make little or no inference. However, prospect theory could be useful precisely for providing quantitative limits or trade-offs to this parameter, similarly to Kim et al. (2010).[23]

More recently, Nielsen & Sebald (2010) have constructed a framework based on DPGs, but completing it with dynamic unawareness.[24] In their model, unawareness impacts players' beliefs, which, in turn, influences their strategic choices.

C. PROSPECT THEORY

Moving on, arguably one of the most important additions to psychological game theory can come from prospect theory. Whereas expected utility has long been the descriptive model for rational choice and economic behaviour, psychologically-based prospect theory has been widely accepted as a more complex alternative.[25] The theory pioneered by Kahneman & Tversky (1979) singles out inconsistencies in expected utility by building on the basis of psychological underpinnings. The paper defines such inconsistencies as the certainty effect, when people putt less weight on outcomes with low probabilities, effect which they prove to

22 See Rabin (1993), Kirchsteiger (1994), Dufwenberg & Kirchsteiger (2004), Falk & Fischbacher (2006) Segal & Sobel (2007), or Dhaene & Bouckaert (2010) for different models of belief-dependent preferences.

23 See the next section for a description of that paper and its results.

24 Following the work of Heifetz et al. (2006, 2008, 2011).

25 See Von Neumann & Morgenstern (1944), Friedman & Savage (1948), or Keeney & Raiffa (1993) for the bases of expected utility theory.

contribute to risk aversion and risk seeking behaviour. Furthermore, they also define the isolation effect as the tendency of forsaking common constituents of different prospects.

Moreover, the theory assigns value to losses and gains rather than final assets, based on the proven assumptions that the human cognitive system intuitively responds to comparisons between values rather than absolutes. As losses and gains are defined as deviations from a baseline or reference level, subsequent research on prospect theory has focused on defining such a neutral reference point, which can vary from one prospect to another based on a number of factors, among which the player's expectations or beliefs, thus providing a natural link to psychological game theory.[26]

Furthermore, prospect theory defines decision weights in order to capture psychological evidence such as people underweighting outcomes with low probabilities. These are subjective preferences applied to each outcome and measure the impact of events on the attractiveness of a prospect, being inferred from observed choices between different prospects. As they apply to prior probabilities of events, they can be an important addition to psychological games, as I will argue in the next section.

In Kószegi & Rabin (2006), a model of reference-dependent preferences is developed by combining reference-dependent gain-loss utility with standard consumption utility, assuming that reference points are represented by recent expectations about outcomes. The model is one of the few which try to develop a comprehensive, general framework and integrate it into economic theory, rather than merely explain experimental results. The next paragraphs describe the model, its assumptions and solution concepts.

Utility is defined in accordance with the properties of the value function of Kahneman & Tversky (1979), and comprises two parts: consumption utility and gain-loss utility. Consumption utility $m(c)$ is the intrinsic property, independent of any reference level, of payoff-based utility based on a K-dimensional consumption bundle c. Gain-loss utility $n(c|r)$, on the other hand, depends on a reference level r. Thus, total utility can be written as $u(c|r) \equiv m(c) + n(c|r)$. Both consumption utility and gain-loss utility are assumed separable across the K dimensions, i.e. $m(c) \equiv \sum_k m_k(c_k)$ and $n(c \mid r) \equiv \sum_k n_k(c_k \mid r_k)$. Gain-loss utility in one dimension depends universally and uniquely on the rapport between consumption utility and the reference point in that dimension, so that $n_k(c_k \mid r_k) \equiv \mu(m_k\left(c_k\right) - m_k(r_k))$. It is thus μ which bares resemblance to the value function mentioned above.

As the model allows for stochastic outcomes and reference points, it assumes that the former are evaluated at the average utility of each outcome compared to each realisation of the reference point. Furthermore, the paper defines the reference point as recent (probabilistic) beliefs about outcomes. The discussion of choice of reference points contrasts beliefs about outcomes with *status quo*-based references, and the paper sets to give a number of examples which contradict the view that gain-loss utility is based on status quo, rather than expectations. One of the examples involves, on the one hand, the (unexpected) monetary gain in a lab

26 See Tversky & Kahneman (1991), Munro & Sugden (2003), Sugden (2003) and Kószegi & Rabin (2006, 2009).

experiment a person might get, which will be assessed as a large gain relative to her status quo. On the other hand, a wage raise which is slightly lower than what an employee expected to get following the company's reported profits will not be seen as a gain relative to the status quo, but as a loss relative to her expectations, even if, say, the rise is 100,000 EUR, extremely large compared to the status quo, which in this situation is the person not receiving it at all.[27] Another example involving psychological rather than monetary payoffs is when a person misses a flight to a holiday destination. The status quo here is the person not being on vacation, so missing the plane would not count as a loss relative to this status quo. However, the person expected to go on vacation, so missing the plane is a loss relative to her expectation of the outcome of her decision to buy the plane ticket.

In order to predict outcomes, the model assumes rational expectations in order to define *personal equilibrium* as the state in which optimal behaviour conditioned by expectations leads to a stochastic outcome that is identical to expectations. Rationality is based on the assumption that people can predict their own behaviour.[28] Extending the framework, the model defines a *preferred personal equilibrium*, which singles out the highest utility personal equilibrium.

As stated before, Kószegi & Rabin (2006) define overall utility as made up of two distinct parts: consumption utility and gain-loss utility, leading to the probabilistic formula $U\left(F \mid G\right) = \int_c \int_r u\left(c|\mathrm{r}\right) dG\left(r\right) dF(c)$, with the assumption that preferences are linear.[29] This formulation of utility captures the fact that gains or losses derived from an outcome are evaluated by comparing that outcome to all possible expected outcomes. Assuming the function is additively separable across the *K* dimensions, utility is assessed independently in each dimension.

Remember that they propose a strong relationship between consumption utility and gain-loss utility, assuming that gain-loss utility in dimension *k* depends in a universal way on the rapport between consumption utility and the reference point in that dimension, so that $n_k(c_k \mid r_k) \equiv \mu(m_k\left(c_k\right) - m_k(r_k))$.

The properties of total utility are consistent with loss aversion and diminishing sensitivity, defined as greater sensitivity to changes closer to the reference level than farther away from it.

These main properties of the universal gain-loss function lead to the following secondary – but equally important – properties:

I. Holding the outcome fixed, lower reference points provide greater utility than higher ones;

[27] The status quo might also be the wage the employee gets before the raise. However, this is an expectation-based status quo, which lies in between the definition of a status quo and expectation-based reference point. Larsen et al. (2004) provide experimental evidence that subjects have mixed emotions (positive and negative simultaneously) due to counterfactual comparisons.

[28] For a discussion of beliefs about own behaviour, see Battigalli & Dufwenberg (2009).

[29] Assumption of linearity is made for ease of description.

II. Assuming linearity of m, u has the same properties as μ. This insures that loss aversion in u is quantitatively less than in μ.

Defining a player's probabilistic beliefs as the distribution Z over his possible choice sets $\{C_v\}_{v\in\mathbb{R}}$, with $C_v \subset \Delta(\mathbb{R}^k)$, a selection $S_v \in C_v$ is a *personal equilibrium* if $U(S_v \mid \int S_v dZ(v)) \geq U(S'_v \mid \int S_v dZ(v))$for all $v \in$ R, $S'_v \in C_v$. This solution concept thus imposes rational expectations as the only restriction to its existence. The definition of personal equilibrium can alternatively be formulated as the maximisation of utility U based on rational expectations about the distribution of outcomes $\int S_v dZ(v)$), which serve as reference point. Transposed to the play path, personal equilibrium becomes Nash equilibrium conditioned by reference-dependent preferences based on lagged expectations.

As the definition allows multiple personal equilibria, a second type of (stronger) solution concept is postulated: a selection $S_v \in C_v$ is a *preferred personal equilibrium* if it is a personal equilibrium and for all $S'_v \in C_v$ we have

$$U\left(\int S_v dZ(v) \mid \int S_v dZ(v)\right) \geq U\left(\int S'_v dZ(v) \mid \int S'_v dZ(v)\right).$$

This second form of equilibrium predicts that players maximize their utility function m in deterministic settings, so the notion captures results of classical game theoretical solution concepts in high-probability environments.

The notions developed in the paper can arguably be incorporated in psychological game theory. However, the authors admit that their model makes some extreme assumptions, such as the strong relationship between consumption and gain-loss utilities or, indeed, that the reference point is fully defined by lagged expectations.

Recent experiments have confirmed prospect theory's predictions. Some results show mathematical patterns that can be adopted to describe effort-based decision making behaviour in a robust manner.[30] Kim et al. (2010) have found these patterns to be characterized by a simple power law, its robustness and rigorous definition suggesting its applicability to game theory. Their set of three experiments involved participants responding by a simple keypress procedure to different stimuli in the form of pictures with potential motivational value: a set of images of faces, a set of affective images and pictures of food. These stimuli were positive and negative, in order to assess whether there is any trade-off between approach and avoidance behaviour. The study revolves around the notion of *relative preference*, defined as the degree with which individuals approach or avoid events based on their features. In economic terms, it can be phrased as the effort an individual is prepared to make in order to avoid an outcome that she assesses as bad or to approach a seemingly positive outcome. In this experiment, the utility loss of an individual is given by the effort made to press a key either to approach or reject an outcome from the sets presented above.

30 Decision making involving relative preference is based on measuring the effort individuals make in order to avoid (perceived) negative outcomes or to seek positive ones.

The results of the study seem to replicate the theory of prospects. The resulting value function describing individuals' approach-avoidance behaviour resembles the function characterised by the five properties in Kahneman & Tversky (1979) and in Kószegi & Rabin (2006).

Although the study admittedly does not involve economic decisions, slight design changes could reveal if the same assumptions hold under economic terms and could be used as (further) confirmation.[31]

III. PROPOSED EXTENSION

My main proposal is that models should incorporate notions of prospect theory into the psychological game theory framework. More specifically, using insights on decision weights, a weighting function for outcomes based on conditional beliefs of the form $\pi(\mu)$ can be envisaged. This function is non-linear and increasing in μ and its slope is a measure of sensitivity of preferences to marginal changes in probabilistic beliefs (Kahneman & Tversky 1979). Decision weights thus complement the DPG framework by bounding conditional beliefs, at the same time keeping the dynamic properties of the updated beliefs system, so that it describes conditional probability systems with the same properties as in DPGs. Consequently, some extra properties seem natural. Firstly, higher-order beliefs about events should have less impact on utility, as intuitively players attribute less importance to beliefs about beliefs about beliefs…of others. The implementation of such concept can be done by changing either the definition of hierarchical beliefs μ, or through the properties of decision weights $\pi(\mu)$ applied to such beliefs.

Furthermore, the definition of utility should allow for a direct dependence on beliefs, but being re-specified as a function f constrained by a set of reference points, deviations from these determining the shape of the function, similar to Kószegi & Rabin (2006). Thus, f is concave for gains (or positive deviations from the reference), and convex for losses (negative deviations from the reference), its slope being greater (in absolute value) under losses than gains, in accordance with prospect theory.

A reference point can be the payoff which a player expects to get by choosing a given strategy, and assumes what the others will do. Thus, beliefs are included in utility as reference points, a point also made in Kószegi & Rabin (2006). Updating of beliefs based on previous play must also be maintained in order to capture the dynamic characteristics of decision making, thus inducing updating of reference points. Consequently, Kószegi & Rabin's (2006) model would be modified in the sense that the reference point is determined by updated expectations, and not by *recent expectations*.

31 KIM ET AL. (2010) argue that the key press procedure might not be too relevant for economic decisions, as it captures a "within-person centric context".

An intuitive next step would be to allow utility to decrease in the number of updates of the reference point a player makes during the game. The reason for this lays in the following example: suppose Joe wants to take Jane out to dinner tonight. His initial belief is that she will say yes. However, when he calls Jane she says she would prefer to go out tomorrow, so John's expectations change. The next day she postpones again for a day later, and so on. Even not supposing this goes on forever, intuition tells us that John will not get the same satisfaction from going out later rather than when he first expected to go out, even if he updates his beliefs during the course of play. Of course, the sequential psychological equilibrium notion in DPGs helps filter this kind of game play out, so that only consistent beliefs lead to equilibrium. However, as the paper admits, one could make the case that predicting game play stands not only on players synchronizing on equilibrium strategies, but also on non-equilibrium play. The notions of personal and preferred personal equilibria can be integrated in this model alongside psychological equilibrium concepts in order to capture the more complex utility function described above.

The implications of such a complex model are as varied as the domains in which game theory is employed. It will be restricted, though, to fields where beliefs play a part in decision making. Nevertheless, given that beliefs are a general human trait, one could argue the importance of this research for most social domains.

I anticipate that issues such as cooperation equilibria under Cournot or Bertrand competition could be investigated in the context of such a framework. Repeated games in which players collaborate to form cartels under the threat of deviation/retaliation are foremost candidates, given their overt psychological characteristics. One should also expect a tangible contribution to topics like bargaining, auctions, voting systems, social network theory, social choice theory, political economy, fair division frameworks, and basically any subject that can be modelled as a game where players (are presumed to) express belief-dependent preferences.

IV. CONCLUSION

Game theory is the standard language for analysing interactions in social sciences. Whether it is behaviour of companies, costumers, employees or plain individuals within society, game theory is used to forecast outcomes of these interactions between different components of human decision making. As traditional models tend to rely on introspection and exogenous parameters, psychological game theory starts from the hypothesis that experimental results show how players actually behave in strategic interactions, and thus expands the mathematical framework of traditional theory.

The many areas of possible use for game theory serve as both an incentive and deterrent for extensions. This paper shows that a relatively intuitive fact such as belief-dependent preferences affecting decision making has been only recently hypothesised and mathematically pinned in economic theory, although other social sciences have been developing such models based on what they considered obvious experimental results advancing such assumptions. On the other hand, those sciences have had trouble capturing hypothesised behaviour with the mathematical rigorousness present in economic models. Only recently have psychologists, neurologists and economists come together with the aim of better understanding underlying human traits and their influence on decision making in a wide variety of social situations.[32]

Psychological game theory, although it extends traditional models up to the point of including belief-dependent behaviour, could arguably profit from methods and assumptions already present in prospect theory. This paper argues that bounding infinite-order beliefs might increase prediction power of current models, while at the same time paving the way towards testing how interactions between players might be used to induce a certain type of desired behaviour, making use of emotional responses.

Assuming that economic behaviour can only be seen as a part of a more complex decision making behaviour specific to humans, then there are no reasons to separate the analysis of psychological constructs and economic strategy, as has been done much too often.

Finally, this paper lays the groundwork for a more ample research agenda, which may include not only building on existing mathematical models and experimental results, but also expanding the existing literature with specific empirical analyses and experiments.

V. REFERENCES

Allais, M., 1953, Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Americaine, *Econometrica: Journal of the Econometric Society*, 503–546.

Battigalli, P. and Dufwenberg, M., 2009, Dynamic psychological games. *Journal of Economic Theory* 144, 1–35.

Battigalli, P. and Dufwenberg, M., 2007, Guilt in games. *The American economic review* 97(2), 170–176.

Battigalli, P. and Dufwenberg, M., 2011, Incorporating Belief-Dependent Motivations in Games, *Discussion Papers*.

Battigalli, P. and Siniscalchi, M., 1999, Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games, *Journal of Economic Theory*, 88(1), 188–230.

Baumeister, R.F., Stillwell, A.M. and Heatherton, T.F., 1994, Guilt: An interpersonal approach, *Psychological bulletin*, 115(2), 243.

Becker, G.S., 1968, Crime and Punishment: An Economic Approach, *The Journal of Political Economy*, 76(2), 169–217.

[32] Chang et al. (2011) is a foremost example of a team of scientists with different specialisations working together in an attempt to reconcile cutting-edge technology with state of the art theoretical models in psychology, neurology and behavioural economics.

Bernheim, B.D., 1994, A theory of conformity, *Journal of Political Economy*, 841–877.
Camerer, C., 2003, *Behavioral game theory: Experiments in strategic interaction*, Princeton University Press.
Caplin, A. and Leahy, J., 2004, The Supply of Information by a Concerned Expert, *Economic Journal*, 487–505.
Chang, L.J. et al., 2011, Triangulating the Neural, Psychological, and Economic Bases of Guilt Aversion, *Neuron*, 70(3), 560–572.
Dhaene, G. and Bouckaert, J., 2010, Sequential reciprocity in two-player, two-stage games: An experimental analysis. *Games and Economic Behavior*, 70(2), 289–303.
Dufwenberg, M., 2002, Marital investments, time consistency and emotions, *Journal of Economic Behavior & Organization*, 48(1), 57–69.
Dufwenberg, M. and Kirchsteiger, G., 2004, A theory of sequential reciprocity, *Games and Economic Behavior*, 47(2), 268–298.
Dufwenberg, M. and Lundholm, M., 2001, Social norms and moral hazard, *The Economic Journal*, 111(473), 506–525.
Elster, J., 1998, Emotions and economic theory, *Journal of economic literature*, 36(1), 47–74.
Erev, I. and Roth, A.E., 1998, Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria, *American economic review*, 848–881.
Falk, A. and Fischbacher, U., 2006, A theory of reciprocity, *Games and Economic Behavior*, 54(2), 293–315.
Forsythe, R. et al., 1994, Fairness in simple bargaining experiments, *Games and Economic behavior*, 6(3), 347–369.
Friedman, M. & Savage, L.J., 1948, The utility analysis of choices involving risk, *The Journal of Political Economy*, 56(4), 279–304.
Geanakoplos, J., Pearce, D. and Stacchetti, E., 1989, Psychological games and sequential rationality, *Games and Economic Behavior*, 1(1), 60–79.
Gilboa, I. and Schmeidler, D., 1988, Information dependent games:: Can common sense be common knowledge?, *Economics Letters*, 27(3), 215–221.
Heifetz, A., Meier, M. and Schipper, B., 2008, A canonical model for interactive unawareness, *Games and Economic Behavior*, 62(1), 304–324.
Heifetz, A., Meier, M. and Schipper, B., 2006, Interactive unawareness, *Journal of Economic Theory*, 130(1), 78–94.
Heifetz, A., Meier, M. and Schipper, B.C., 2011, Dynamic unawareness and rationalizable behavior, *Discussion Papers*.
Huck, S., Normann, H.T. and Oechssler, J., 1999, Learning in Cournot oligopoly – An experiment, *The Economic Journal*, 109(454), 80–95.
Kahneman, D. and Tversky, A., 1979, Prospect theory: An analysis of decision under risk, *Econometrica: Journal of the Econometric Society*, 263–291.

Keeney, R.L. and Raiffa, H., 1993, *Decisions with multiple objectives: Preferences and value tradeoffs*, Cambridge University Press.

Kim, B.W. et al., 2010, Recurrent, Robust and Scalable Patterns Underlie Human Approach and Avoidance, J. Lauwereyns (ed.), *PLoS ONE*, 5, p.e10613.

Kirchsteiger, G., 1994, The role of envy in ultimatum games, *Journal of economic behavior & organization*, 25(3), 373–389.

Kolpin, V., 1992, Equilibrium refinement in psychological games, *Games and Economic Behavior*, 4(2), 218–231.

Kószegi, B., 2006, Emotional Agency, *Quarterly Journal of Economics*, 121(1), 121–155.

Kószegi, B. and Rabin, M., 2006, A model of reference-dependent preferences, *The Quarterly Journal of Economics*, 121(4), 1133.

Kószegi, B. and Rabin, M., 2009, Reference-Dependent Consumption Plans, *American Economic Review*, 99(3), 909–936.

Kreps, D.M. and Wilson, R., 1982, Sequential equilibria, *Econometrica: Journal of the Econometric Society*, 863–894.

Larsen, J.T. et al., 2004, The agony of victory and thrill of defeat, *Psychological science*, 15(5), 325.

Loewenstein, G.F. et al., 2001, Risk as feelings, *Psychological bulletin*, 127(2), 267.

Munro, A. and Sugden, R., 2003, On the theory of reference-dependent preferences, *Journal of Economic Behavior & Organization*, 50(4), 407–428.

Von Neumann, J. and Morgenstern, O., 1944, *Theory of games and economic behavior*, Princeton, NJ: Princeton University Press.

Nielsen, C.S. and Sebald, A., 2010, Unawareness in Dynamic Psychological Games, *Discussion Papers*.

Offerman, T., Potters, J. and Sonnemans, J., 2002, Imitation and belief learning in an oligopoly experiment, *Review of Economic Studies*, 69(4), 973–997.

Rabin, M., 1993, Incorporating fairness into game theory and economics, *The American Economic Review*, 1281–1302.

Roth, A.E. and Erev, I., 1995, Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term, *Games and Economic Behavior*, 8(1), 164–212.

Segal, U. and Sobel, J., 2007, Tit for tat: Foundations of preferences for reciprocity in strategic settings, *Journal of Economic Theory*, 136(1), 197–216.

Simon, H.A., 1955, A behavioral model of rational choice, *The quarterly journal of economics*, 69(1), 99.

Sugden, R., 2003, Reference-dependent subjective expected utility, *Journal of economic theory*, 111(2), 172–191.

Thaler, R., 1980, Toward a positive theory of consumer choice, *Journal of Economic Behavior & Organization*, 1(1), 39–60.

Tversky, A. and Kahneman, D., 1974, Judgment under uncertainty: Heuristics and biases, *Science*, 185(4157), 1124.

Tversky, A. and Kahneman, D., 1991, Loss aversion in riskless choice: A reference-dependent model, *The Quarterly Journal of Economics*, 106(4), 1039.

Vega-Redondo, F., 1997, The evolution of Walrasian behavior, *Econometrica: Journal of the Econometric Society*, 375–384.

Young, H.P., 1993, The evolution of conventions, *Econometrica: Journal of the Econometric Society*, 57–84.