Evolution of Theories of Mind*

Erik Mohlin†
University College London

May 22, 2011

Abstract

This paper studies the evolution of peoples' models of how other people think – their theories of mind. First, this is formalized within the level-k model, which postulates a hierarchy of types, such that type k plays a k times iterated best response to the uniform distribution. It is found that, under plausible conditions, lower types co-exist with higher types. The results are extended to a model of learning, in which type k plays a k times iterated best response the average of past play. The model is also extended to allow for partial observability of the opponent's type.

Keywords: Theory of Mind; Depth of Reasoning; Evolution; Learning; Level-k; Fictitious Play; Cognitive Hierarchy.

JEL codes: C72, C73, D01, D03, D83.

*Valuable comments were provided by Ola Andersson, Vincent Crawford, Stefano Demichelis, Tore Ellingsen, Drew Fudenberg, Robert Östling, and Bill Sandholm. A special thank goes to Jörgen Weibull for advice and encouragement. The paper has also benefitted from comments made by audiences at the European Winter Meeting of the Econometric Society, Budapest 2009, the 1st UECE Lisbon Meeting: Game Theory and Applications 2009, the 4th Nordic Workshop in Behavioral and Experimental Economics, Oslo 2009, as well as seminars at the University of Bonn (Neuroeconomics Lab), the University of Amsterdam (CREED), University College London, and the Stockholm School of Economics. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.

†E-mail: e.mohlin@ucl.ac.uk. Mail: Department of Economics, University College London, Gower Street, London WC1E 6BT, United Kingdom. Telephone: +44 (0)20 7679 5485. Fax: +44 (0)207 916 2775.

1 Introduction

In order to decide what strategy to choose, a player usually needs to form beliefs about what other players will do. This requires the player to have a model of how other people form beliefs – what psychologists call a *theory of mind* (Premack and Wodruff (1979)). In this paper I study the evolution of players' theories of mind, both in the form of their models of how other players form initial beliefs, and in the form of their models of how other players learn.

When people play a game for the first time, their initial behavior rarely conforms to a Nash equilibrium.[1] In such situations, behavior is often more successfully predicted by the level-k model (Stahl and Wilson (1995) and Nagel (1995)), and the related cognitive hierarchy (Camerer et al. (2004)), and noisy introspection models (Goeree and Holt (2004)).[2] According to these models, people display limited depth of reasoning, when forming beliefs about other peoples' behavior. Moreover, people differ with respect to how they form beliefs. The heterogeneity is represented by a set of cognitive types, such that higher types form more sophisticated beliefs. According to the level-k model, the lowest type, type zero, does not form any beliefs and randomizes uniformly over the strategy space. An individual of type $k \geq 1$ believes that everyone else belongs to type $k-1$, and hence plays a k times iterated best response to the uniform distribution. The model is often specified so that type 0 merely exists in the minds of higher types (Costa-Gomes and Crawford (2006)).

In order to study evolution of theories of mind in the context of learning, I consider an extension of fictitious play. According to fictitious play all individuals best respond to the average of past play. If some individuals follow this rule, it is natural to hypothesize that some more sophisticated individuals realize this, and play a best response to the best response to the average of past play. And it seems quite possible that some individuals think yet another step and play a twice iterated best response to the average of past play. Continuing in this way one arrives at a hierarchy of types, using increasingly complex models of how other people learn. I refer to the resulting model as *heterogeneous fictitious play.* A related model is

[1] See Camerer (2003). This claim presupposes some assumption about preferences. In the mentioned studies, the preferences needed to make the observed behavior conform to Nash equilibrium seem like a much less reasonable explanation than attributing the behavior to some form of incorrect expectations. See the discussion in Costa-Gomes and Crawford (2006).

[2] For experimental evidence on this see Camerer et al. (2004), Costa-Gomes and Crawford (2006), and Camerer (2003). Coricelli and Nagel (2009) present neuroeconomic evidence. The models have also been applied to e.g. auctions (Crawford and Iriberri (2007)), communication (Crawford (2003), Ellingsen and Östling (2010)), and marketing decisions (Brown et al. (2008)).

proposed by Stahl (1999), and subjected to experimental testing by Stahl (2000).[3]

Many games with important consequences are only played a few times during a life-time, with limited scope for learning. For example, the choice of a career and the choice of a mate are parts of complicated games which most people play only once, or a few times. Still, the strategic thinking employed in such interactions should be immensely important for the success of a person, in economic as well as biological terms. Other games are played many times, with feedback that allows the players to learn. The evolutionary advantage of an accurate model of how other people learn should be obvious in such cases.

The contribution of this paper is to provide an evolutionary analysis of initial responses, as described by the level-k model, and modes of learning, as formalized by the heterogeneous fictitious play model. In particular I explain why evolution may lead to a state where people display heterogeneous and limited depths of reasoning.[4] There are only a few studies of evolution of cognitive types which can be interpreted as being about initial responses; Stahl (1993), Banerjee and Weibull (1995), Stennek (2000), and Samuelson (2001*a*). These models differ in important respects from the model put forward in this paper, and they do not identify the same mechanisms shaping the distribution of types (see section 5). There is a vast literature studying properties of different formal learning rules (see e.g. Fudenberg and Levine (1998), and Sandholm (2011)), but there are hardly any studies of the evolutionary properties of the learning rules themselves. Exceptions include Heller (2004) and Josephson (2008).

Since evolution acts upon types rather than strategies, it is possible to study evolution across different games, assuming that individuals are matched to play games that are drawn from a class of games. In contrast, most of the literature on evolution in games focuses on one game at a time, and among those who study evolution across games, most papers only consider games which have identical strategy spaces; to my knowledge Samuelson (2001*a*) is the only exception.[5]

The level-k model, implicitly assumes that players lack specific information about the cognitive types of their opponents. I extend the level-k model to allow types to be partially observed. Such an extension seems essential in order to

[3] Fudenberg and Takahashi (2011) model fictitious play with heterogeneous beliefs.

[4] One might object that the heterogeneity could be due to random variation. However, there is evidence that strategic reasoning is implemented by specialized modules in the brain (Cosmides and Tooby 1992), and it has been argued (Penke et al. 2007) that variation in such traits is best explained by frequency dependent selection, rather than by random variation.

[5] Heller (2004), Mengel (2009), and Steiner and Stewart (2008) model evolution and learning across game with identical strategy spaces. Haruvy and Stahl (2009) experimentally investigate learning across games with different strategy spaces.

capture situations where an unfamiliar game is played by individuals who have some information about their opponents' ways of thinking. In the kind of small-scale societies that characterized much of our evolutionary past, such information should have been common. To model this I assume that higher types recognize and best respond to lower types, but that lower types do not know how the higher types think. These assumptions are motivated by the way types are defined: The cognitive type of an individual represents that individual's ability to understand how other people think. Thus, being of a high type means that one is good at understanding how other people think, i.e. that one is good at detecting their type.[6]

From an evolutionary perspective, the potential advantage of having a better theory of mind has to be weighed against the cost of increased reasoning capacity. I abstract from such costs in the formal analysis, but note that they limit the survival chances of higher types (see section 5.3).

For the evolutionary analysis of the level-k model, consider a large population of individuals of different types, who are randomly matched to play a symmetric two-player game. The types of the matched pair of individuals, determine their actions, and hence their payoffs. The population fractions of types evolve in proportion to their average payoffs, in accordance with the replicator dynamic. In the case of the heterogeneous fictitious play model, one must consider both learning and evolution. In each period, individuals are drawn to play one game many times, each time with a different opponent. The average payoff over these interactions constitutes the evolutionarily relevant payoff in the current period.

The results presented in this paper identify conditions under which evolution leads to – or does not lead to – states where the highest type, and types behaving like the highest type, dominate. I restrict attention to finite symmetric two-player games in which all level-k types, except type 0, play pure strategies. First I consider the standard level-k model (without any observability of types). It turns out that there is an important distinction to be made, between type-acyclic and type-cyclic games. A game is *type-acyclic* game if there exists a finite k, such that the strategy that is a k times iterated best response against the uniform distribution, is also a best response to itself. Games that are not type-acyclic are *type-cyclic*. Provided that the set of types is large enough relative to the number of strategies, the set of states in which only the types behaving like the highest type exist, is asymptotically stable if and only if the game is type-acyclic. Still, even in type-acyclic games (e.g.

[6] In the literature on preference evolution it is well known, that if players have complete information about preferences, then evolution may favor preferences that do not coincide with material payoffs (Dekel et al. (2007)). As pointed out e.g. by Samuelson (2001*b*), it is important that assumptions about observability are well motivated and not ad hoc. See also section 5.3.

in some dominance solvable games) evolution may lead to asymptotically stable sets, in which the highest type, and all types behaving like the highest type, are extinct. There is a sufficient condition, dubbed weak best reply dominance (satisfied e.g. by all strictly supermodular games), that guarantees the convergence to a state where everyone behaves like the highest type. In type-cyclic games different results obtain. Consider stable games with an interior evolutionarily stable strategy (ESS), such as the Hawk Dove game. In such games it can be shown that evolution leads to an asymptotically stable set of states, in which types that do not behave like the highest type, co-exist with types that do behave like the highest type. Under some circumstances, all types may co-exist. The intuition for these results is that there is an incentive not to behave, and hence not to think, like the opponent in these games. Next, suppose that individuals play games from a *set of games* consisting of a type-acyclic game and a type-cyclic game. If the set of types is large enough relative to the number of strategies, then the set of states where everyone behaves like the highest type, is unstable. For the special case of a set of games consisting of a 2-strategy coordination game, a 2-strategy stable game, and a Travelers Dilemma game, it is shown that evolution may converge to a state with heterogeneous types, starting from any initial condition.

The results for the heterogeneous fictitious play model are similar to those of for the standard level-k model. In games satisfying weak best reply dominance, evolution leads to states where only the highest types survive. In the game of Shapley (1964) ordinary fictitious play cycles, but evolution according to the heterogeneous fictitious play model converges to a state where different types co-exist, such that behavior corresponds to the Nash equilibrium in all periods.

Introducing partial observability of types into the level-k model, alters the results. In games that satisfy weak best reply dominance, exemplified by the Travelers' Dilemma, evolution may lead to the co-existence of different types, who do not behave like the highest type. This happens because lower types choose strategies which induce more efficient outcomes than what is obtained when two high types play. In some stable games adding partial observability improves the prospects for different types to co-exist, but in other stable games the effect is reversed.

The rest of the paper is organized as follows: The next section presents the main model, focusing on the level-k model. The results for the level-k model are presented in section 3. Section 4 contains the heterogeneous fictitious play model and the the level-k model with partial observability. Section 5 discusses related literature and other issues. Section 6 concludes. Proofs can be found in the appendix. Additional proofs and results are relegated to the supplementary material.

2 Model

2.1 Preliminaries

Consider a symmetric two-player normal form game G with a finite pure strategy set S and mixed strategy set $\Delta(S)$. Payoffs are given by $\pi : S \times S \to \mathbb{R}$, where $\pi(s, s')$ is the payoff to a player using strategy s against strategy s'. For mixed strategies the expected payoffs are given by $\tilde{\pi} : \Delta(S) \times \Delta(S) \to \mathbb{R}$ where $\tilde{\pi}(\sigma, \sigma')$ is the payoff to a player, using strategy σ against strategy σ'. With slight abuse of notation let s denote the degenerate mixed strategy that puts all weight on pure strategy s. Let $\beta : \Delta(S) \twoheadrightarrow S$ be the pure best reply correspondence. If the best response is unique I write $\beta(\sigma) = s$ rather than $\beta(\sigma) = \{s\}$. The uniform randomization over the set of pure best responses to σ, is denoted $\bar{\beta}(\sigma)$. Again, with slight abuse of notation, the expression $\beta(s)$ stands for the pure best response to the mixed strategy that puts all weight on the pure strategy s (and similarly for $\bar{\beta}$).

A population consists of a finite set of cognitive types $K = \{0, 1, 2, ..., \kappa\}$. The set of probability distributions over K is $\Delta(K)$, so a *population state* is a point

$$x = (x_0, x_1, ...x_\kappa) \in \Delta(K).$$

(For reasons that will be explained below, some of the analysis disregards type 0, so that the set of types is $K \backslash \{0\}$, and a population state is a point $x \in \Delta(K \backslash \{0\})$.) Suppose that an individual of type k and an individual of type k' play a symmetric two-player normal form game G. Let $\sigma(k) \in \Delta(S)$ be the strategy that an individual of type k plays. For a given game G, the expected payoff of type k, against type k', is $\tilde{\pi}(\sigma(k), \sigma(k'))$, so the expected payoff of type k, in state x, is given by the function $\Pi_k^G : \Delta(K) \to \mathbb{R}$, with

$$\Pi_k^G(x) = \sum_{k' \in K} x_{k'} \tilde{\pi}(\sigma(k), \sigma(k')). \tag{1}$$

Suppose (in order to examine evolution across games) that individuals are randomly matched to play a game which is drawn from a finite set of games $\mathcal{G}$, according to a probability measure μ. The individuals are informed about the game that has been drawn. The expected payoff of type k, in state x, is now given by the function $\Pi_k^{\mathcal{G}} : \Delta(K) \to \mathbb{R}$, with

$$\Pi_k^{\mathcal{G}}(x) = \sum_{G \in \mathcal{G}} \mu^G \Pi_k^G(x),$$

where μ^G is the probability of game G.

2.2 The Level-k Model

According to the level-k model, type 0 randomizes uniformly over the strategy space.[7] Each type $k \geq 1$ best replies to type $k-1$. Let U denote the uniform distribution over S, and let $\beta^i(U)$ denote the $i \geq 0$ times iterated best response to the uniform distribution, with the convention $\beta^0(U) = U$. In order to avoid unnecessary complication I will – throughout the whole paper – restrict attention to symmetric *games where the best responses of all types are unique*, i.e. $\beta^k(U)$ is a singleton for all $k \geq 0$.[8] Thus, type $k \in K$ plays $\sigma(k) = \beta^k(U)$. If the payoffs associated with different strategy profiles are drawn independently, according to some continuous measure defined over an interval on the real line, then the uniqueness assumptions is satisfied for all but a measure zero set of games.

2.3 Evolution

For a given game G, the average payoff in the population, in state x, is

$$\bar{\Pi}^G(x) = \sum_{k=0}^{\kappa} x_k \Pi_k^G(x).$$

Evolution of types is determined by the *replicator dynamic,* for all $k \in K$,

$$\dot{x}_k = [\Pi_k^G(x) - \bar{\Pi}^G(x)]x_k. \tag{2}$$

This defines a vector field $\varphi : \mathbb{R}^{\kappa+1} \to \mathbb{R}^{\kappa+1}$, such that $\dot{x} = \varphi(x)$. Similarly, if the games are drawn from $\mathcal{G}$ according to μ, the average payoff in the population, in state x, is denoted $\bar{\Pi}^{\mathcal{G}}(x)$ and the replicator is defined as above, with $\mathcal{G}$ instead of G.

In the level-k model, each type's behavior is constant across states. It is then easy to verify that the payoffs to the different types, and hence the vector field, is Lipschitz continuous. Let $T \subseteq \mathbb{R}$ be an open interval containing $t = 0$. By the Picard-Lindelöf theorem the system has a unique (local) solution $\xi(\cdot, x^0) : T \to \Delta(K)$ through any initial condition x^0, such that $\xi(0, x^0) = x^0$ and

$$\frac{\partial}{\partial t}\left(\xi\left(t, x^0\right)\right) = [\Pi^G\left(\xi\left(t, x^0\right)\right) - \bar{\Pi}^G\left(\xi\left(t, x^0\right)\right)]\xi\left(t, x^0\right),$$

[7] Other assumptions are possible but I will stick with the assumption that is dominant in the existing literature.

[8] Alternatively one could assume that if the pure best reply is not unique, then the individual follows the principle of insufficient reason and randomizes uniformly over the set of pure best replies. This would complicate the analysis substantially.

for all t. Moreover, since $\Delta(K)$ is compact and the system never leaves $\Delta(K)$, T can be taken to be the whole of $\mathbb{R}$, so that the solution is global.

The notion of a type game will be useful in the analysis below.

Definition 1 *Consider a symmetric two-player normal form game G and a set of types K. The corresponding* ***type game*** *is a symmetric two-player game, where each player's strategy space is K, and the payoff to type-strategy k, against type-strategy k', is $w(k,k') = \tilde{\pi}(\sigma(k), \sigma(k'))$.*

The definition of a type game allows us to apply results from standard evolutionary game theory, where evolution acts upon strategies, to the present setting where evolution acts upon the cognitive types. For instance, if a strategy in the type game is an evolutionarily stable strategy (ESS), then we know that the corresponding state x is asymptotically stable under the replicator dynamic.[9] In the basic level-k model each type plays the same strategy in all states, and all types except type 0 play a pure strategy. Thus, if we disregard type 0 and restrict attention to types $K\backslash\{0\}$, studying evolution of types according to the replicator dynamic is formally the same as studying the replicator on the subset of pure strategies picked by the types in $K\backslash\{0\}$, except for the minor complication that the same pure strategy might be used by more than one type. Partly for this reason, some results below will only be proved for the restricted set of types $K\backslash\{0\}$. Another reason for excluding type 0 from the analysis is that the level-k model is usually estimated under the assumption that type 0 only exists in the minds of higher types. The simple relationship between strategy-evolution in the underlying game and type-evolution in the type game does not hold for the level-k model with partial observability, or the heterogeneous fictitious play model considered below.

3 Results

We are interested in finding stability and convergence properties of, on the one hand, (i) states where only types behaving like the highest type exist, and on the other hand, (ii) states where lower types co-exist with higher types, not playing the same strategies. For any game let $\tilde{K}$ be the *set of types that choose the same strategy as the highest type*;

$$\tilde{K} = \{k \in K : \beta^k(U) = \beta^\kappa(U)\}, \tag{3}$$

[9] Some definitions of stability concepts are provided in the supplement.

and define $\tilde{X}$ to be the *set of states where only types behaving like the highest type exist*;

$$\tilde{X} = \{x \in \Delta(K) : \sum_{k \in \tilde{K}} x_k = 1\}. \tag{4}$$

To address questions regarding the asymptotic stability of $\tilde{X}$, define:

Definition 2 *A game is* ***type-acyclic*** *if there is some $k \in \mathbb{N}$ such that $\beta^k(U) = \beta^{k+1}(U)$. In a type-acyclic game k^{NE} is the smallest number k such that $\beta^k(U) = \beta^{k+1}(U)$. A game is* ***type-cyclic*** *if it is not type-acyclic. In a type-cyclic game k^c is the smallest number k such that $\beta^k(U) = \beta^{k'}(U)$ for some $k' < k$.*[10]

In type-acyclic games k^{NE} is the lowest type that plays a symmetric Nash equilibrium strategy. The strategy played by type k^{NE} will be played by all higher types (since we assume that $\beta^k(U)$ is a singleton for all k). In type-cyclic games k^c is the lowest type k such that the sequence $\{\beta^i(U)\}_{i=1}^k$ contains a cycle. Clearly we always have $k^{NE}, k^c < n$.

Proposition 1 *Suppose that an underlying game is played by level-k types. If the underlying game is type-acyclic and $\kappa \geq k^{NE}$ then the set $\tilde{X}$ is asymptotically stable. If the underlying game is type-cyclic and $\kappa \geq k^c$ then the set $\tilde{X}$ is not asymptotically stable.*

It follows that if κ is large enough relative to the number of strategies n, so that both $\kappa \geq k^{NE}$ and $\kappa \geq k^c$ hold, then $\tilde{X}$ is asymptotically stable *if and only if* the underlying game is type-acyclic. Note that $\kappa > n$ is sufficiently large.

3.1 Type-Acyclic Games

Proposition 1 does not claim that $\tilde{X}$ is the unique asymptotically stable set when the underlying game is type-acyclic and $\kappa \geq k^{NE}$. Neither does it claim that evolution will always end up in that set $\tilde{X}$ under these conditions. For example, type-acyclicity is satisfied by dominance solvable games (supplement, lemma S1), but there are dominance solvable games which generate type games where $\tilde{X}$ is not the only asymptotically stable set; there might be asymptotically stable states where $x_\kappa = 0$ (supplement, example S1). The reason is that the type game generated by an underlying dominance solvable game need not itself be dominance solvable, even when the type space is restricted to $K \backslash \{0\}$. To obtain a condition under which

[10] Young (1993) defines an acyclic game as a game where, for every $s \in S$, there is some (finite) k such that $\beta^k(s) \subseteq \beta^{k-1}(s)$.

$\tilde{X}$ is the unique asymptotically stable set with the whole interior as its basin of attraction, define the property of weak best reply dominance:

Definition 3 *A game satisfies* ***weak best reply dominance (WBRD)*** *if, for all* $k \geq 0$ *and all* s*, it holds that* $\pi\left(\beta(s), \beta^k(s)\right) \geq \pi\left(s, \beta^k(s)\right)$.

The WBRD-property is satisfied by strictly supermodular games (supplement, lemma S2). A game that is not supermodular, but satisfies the WBRD property is the Travelers' Dilemma (Basu (1994)). The two-player symmetric *Travelers' Dilemma* has strategy space $S = \{1, 2, ..., n\}$. The payoff to a player choosing strategy s against strategy s' is

$$\pi(s, s') = \begin{cases} s & \text{if } s = s' \\ s + R & \text{if } s < s' \\ s' - P & \text{if } s > s' \end{cases},$$

for some real numbers $R > 1, P > 0$. Assume $R + P \notin \mathbb{N}$ so that $\beta(s)$ is single-valued for all $s \in S$. In the Travelers' Dilemma there is always an incentive to undercut the opponents choice, by picking a strategy that is one step below the opponent's strategy one obtains a net reward of $R - 1 > 0$. Therefore this game constitutes a social dilemma, in the sense that both players would earn more if they were are able to cooperate and play a high strategy, than if they play the Nash equilibrium profile $(1, 1)$.

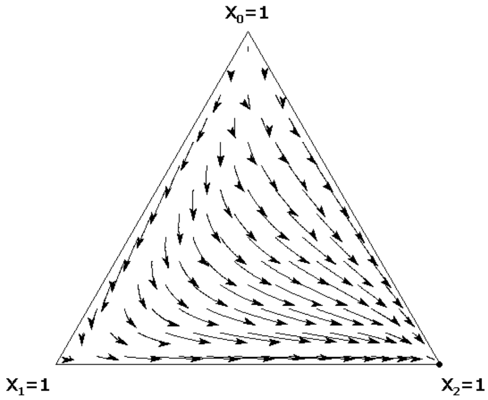
It can be verified that a game satisfying WBRD is type-acyclic (appendix, lemma 1). The following proposition provides a condition for convergence to the asymptotically stable set where only types that behave like the highest type exist.

Proposition 2 *Suppose that an underlying WBRD-game is played by level-k types* $K \backslash \{0\}$*. For any* κ*, the set* $\tilde{X}$ *is the unique asymptotically stable set, with the whole interior as its basin of attraction.*

It is necessary to exclude type 0 in order to obtain the above result. There are games satisfying WBRD, even strictly supermodular games, such that if type 0 is included in the type space, then there is an asymptotically stable set where $x_\kappa = 0$ (supplement, example S2).[11] However, if type 0 asymptotically becomes extinct in a WBRD game, then clearly the above proposition implies that $\tilde{X}$ is the unique

[11] Any game generated by a subset of pure strategies from a supermodular game is itself supermodular. But the strategy that type 0 uses is not a pure strategy of the underlying game. Therefore the cognitive game need not be supermodular when type 0 is included.

Figure 1 *Travelers' Dilemma played by level-k types.* Parameters $R = 3/2$, $P = 1/3$, and $n = 6$. Type space $K = \{0, 1, 2\}$. The state with $x_2 = 1$ (black dot) is asymptotically stable and has the whole interior as its basin of attraction.



asymptotically stable set, with the whole interior as its basin of attraction.[12] Figure 1 illustrates this possibility, for the Travelers' Dilemma with parameters $R = 3/2$, $P = 1/3$, and $n = 6$, played by types $K = \{0, 1, 2\}$.

Even in WBRD-games, the asymptotically stable set of types $\tilde{X}$ can be large. Consider coordination games:

Definition 4 *A game is a* ***coordination game*** *if* $\beta(s) = s$ *for all* $s \in S$.

Coordination games satisfy the WBRD-condition with equality. The simplest example would be a 2×2 coordination game, such as the Stag Hunt game, whose payoff matrix is strategically equivalent to

$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}, \tag{5}$$

for some $a > 1$.[13] Since each pure strategy is the unique best response to itself, all types $k \geq 1$ behave in the same way i.e. $\tilde{X} = \{x \in \Delta(K) : x_0 = 0\}$. Consequently all types $k \geq 1$ earn the same in all states, and all types $k \geq 1$ earn more than type 0. Thus the set of all states where type 0 is absent is asymptotically stable.

Corollary 1 *Suppose that an underlying coordination game is played by level-k types. For any* κ, *the set* $\tilde{X} = \{x \in \Delta(K) : x_0 = 0\}$ *is asymptotically stable and has the whole interior as basin of attraction.*

[12] All phase diagrams were created with *Dynamo* (Sandholm and Dokumaci (2007)).

[13] Two games are strategically equivalent if they share the same best replies and dominance relations. Subtracting scalars from columns and multiplying All generic symmetric 2×2 games fall into one of three categories of strategically equivalent games, (i) coordination games, (ii) games with a unique interior ESS and (iii) dominance solvable games (see Weibull (1995), chapter 1).

3.2 Type-Cyclic Games

Once we move away from the class of type-acyclic games the survival prospects of lower types improve further. A normal form game with payoff matrix $\mathbf{A}$ is said to be *stable* if $\mathbf{A}$ is negative definite with respect to the tangent space.

Definition 5 *A normal form game with an $n \times n$ payoff matrix $\mathbf{A}$ is* ***stable*** *if $v \cdot \mathbf{A}v < 0$ for all $v \in \mathbb{R}^n_0 = \{v \in \mathbb{R}^n : \sum v_i = 0\}$, $v \neq \mathbf{0}$.*

A stable game has a unique ESS. It is easy to see that if the unique ESS is interior then the game is type-cyclic. The simplest example of a stable game with an interior ESS is the *Hawk Dove* game, whose payoff matrix is strategically equivalent to

$$\begin{pmatrix} -b & 0 \\ 0 & -1 \end{pmatrix}, \tag{6}$$

for some $b > 1$.

If all types play different strategies, and if we restrict attention to $K \backslash \{0\}$, then the type game based on an underlying stable game is itself a stable game (appendix, lemma 3), and the replicator dynamic converges to the ESS in stable games (Hofbauer and Sigmund (1988), Sandholm (2011), chapter 6). However, the ESS of the type game will generally not correspond to the same behavior as the ESS of the underlying game. It will be useful to add some more structure that allows us infer what strategies different types choose. In the Hawk-Dove game each strategy is the unique best response to the other strategy. We can generalize this property.

Definition 6 *An n-strategy game is* ***completely cyclic*** *if the strategies can be ordered $s^1, ..., s^n$ such that $\beta(s^i) = s^{i+1_{\mathrm{mod}\, n}}$ for all $i \in \{1, .., n\}$.*

In other words, for all $i \in \{1, .., n-1\}$ the best response to strategy s^i is strategy s^{i+1} and the best response to strategy s^n is strategy s^1. Hence, type i behaves like type i modulo n (abbreviated $i_{\mathrm{mod}\, n}$). Clearly, a completely cyclic game is type-cyclic, and a stable completely cyclic game has a unique interior ESS. Let $\sigma(K, x)$ denote aggregate play at state $x \in \Delta(K)$. For any stable game, define $X^{ESS}(K)$ to be *the set of states where aggregate behavior of types K corresponds to the unique ESS σ^{ESS}*;

$$X^{ESS}(K) = \left\{x \in \Delta(K) : \sigma(K, x) = \sigma^{ESS}\right\}. \tag{7}$$

For the set of types $K \backslash \{0\}$, the set of states $X^{ESS}(K \backslash \{0\})$ is defined similarly. We are now in a position to prove the following result regarding stable games:

Proposition 3 *Suppose that an underlying stable game with an interior ESS, σ^{ESS}, is played by level-k types $K\backslash\{0\}$. For any κ, evolution from any interior initial condition converges to a unique asymptotically stable set $X^* \subseteq \Delta(K\backslash\{0\})$.*

1. *If $\kappa \geq k^c$ then $X^* \cap \tilde{X} = \emptyset$.*
2. *If the game is completely cyclic and $\kappa \geq n$ then $X^* = X^{ESS}(K\backslash\{0\})$ and X^* contains interior states.*
3. *If the game is completely cyclic and $\kappa = n$ then $|X^*| = 1$.*

Part 1 of the above proposition states that if κ is large enough, so that the behavior of the different types $\{\beta^k(U)\}_{k=1}^{\kappa}$ contains a cycle, then evolution converges to some asymptotically stable set, in which not only those types are present, which behave like the highest type. Furthermore, if the underlying game is not only stable, but also completely cyclic, and if there are at least as many types as strategies, then every strategy is played by some type. This means that the set of states, $X^{ESS}(K\backslash\{0\})$, in which aggregate behavior corresponds to the ESS of the underlying game, is non-empty. Under these conditions, part 2 states that $X^{ESS}(K\backslash\{0\})$ is the unique asymptotically stable set, and that it has the whole interior as basin of attraction. Part 3 considers the special case when the number of types is equal to the number of strategies: Since each strategy is chosen by precisely one type, the unique asymptotically stable set $X^{ESS}(K\backslash\{0\})$ now consists of a single state in which all types $K\backslash\{0\}$ co-exist. The general intuition for these results is that the payoffs in stable games and completely cyclic games are such that it is beneficial not to think, and behave, like everyone else. As aggregate behavior approaches the ESS, the payoffs to different strategies are equalized, so that different types, playing different strategies, may earn the same.

The above proposition ignores type 0. In order to prove results for the full set of types K another concept needs to be introduced. The Hawk-Dove game, like the 2×2 coordination game, is a potential game (Monderer and Shapley (1996)).

Definition 7 *An n-strategy game with payoff matrix $\mathbf{A}$ is a **potential game** if $\mathbf{A} = \mathbf{C} + \mathbf{1}\mathbf{r}'$, for some symmetric matrix $\mathbf{C}$ and some column vector $\mathbf{r} \in \mathbb{R}^n$, with transpose $\mathbf{r}'$, and where $\mathbf{1}$ is the column vector with all entries equal to one.*

The type game based on an underlying potential game is itself a potential game, even when we consider the whole set of types K (appendix, lemma 4), and in potential games the replicator dynamic always converges (Sandholm (2001)). However, without additional assumptions it is not possible to tell what strategies the different types will play, and which types that will survive.[14] Therefore the second part of the following proposition is restricted to 2×2-games.

Proposition 4 *Suppose that an underlying stable potential game with an interior ESS, is played by level-k types K. For any κ, evolution from any interior initial condition converges to some (not necessarily unique) asymptotically stable set $X^* \subseteq \Delta(K)$.*

1. *If $\kappa \geq k^c$ then $X^* \cap \tilde{X} = \emptyset$.*
2. *If the game is a 2×2-game then, for any $\kappa \geq 1$, evolution from any interior initial condition converges to the unique asymptotically stable state $X^* = X^{ESS}(K)$.*[15]

The intuition behind proposition 4 is the same as for proposition 3; the payoffs in stable games provide incentives for choosing something else than what the opponent chooses. Again, part 1 of the above proposition states that if κ is large enough, so that the behavior of the different types contains a cycle, then evolution converges to some asymptotically stable set, in which not everyone behaves like the highest type. Part 2 states that for the special case of 2×2-games, evolution leads to the set $X^{ESS}(K)$. Since proposition 4 considers the whole set of types K, including type 0, the set $X^{ESS}(K\})$ has a slightly different structure than the set $X^{ESS}(K\backslash\{0\})$ in proposition 3. In particular, suppose that the ESS fractions of the Hawk and Dove strategies are x_H^{ESS} and x_D^{ESS}, respectively. For any fraction of type zero that satisfies $x_0 \leq 2\min\{x_H, x_D\}$, there is a state in $X^{ESS}(K)$ with this fraction of type 0. Figure 2, illustrates part 2 for a 2×2-game with $b = 2$, played by types $K = \{0, 1, 2\}$.

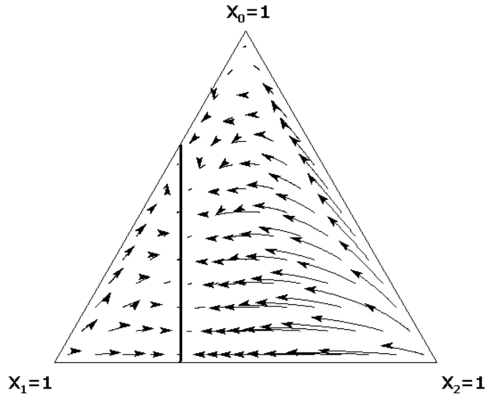
3.3 Evolution across Games

The analysis so far has dealt with one game at a time, resulting in disparate conclusions about the survival of types that do not behave like the highest type. On the one extreme we have the WBRD-games, in which evolution leads to an asymptotically set $\tilde{X}$ where all types not behaving like the highest type are extinct (at

[14]The class of convex potential games includes coordination games, which fall under corollary 1. The class of concave potential games falls into the class of stable games which are handled by proposition 4.

[15]The finding is similar to a result in Banerjee and Weibull (1995). They find that an optimizing type may earn less than a preprogrammed type in games with a unique interior ESS. However, their result only holds for the case of observable types, whereas we have found that lower types may survive even when types are not observable.

Figure 2 *Hawk Dove played by level-k types.* Parameter $b = 2$, and type space $K = \{0, 1, 2\}$. The set $X^* = X^{ESS}(K)$ (thick line) is asymptotically stable and has the whole interior as its basin of attraction.



least if we ignore type 0). On the other extreme we have the completely cyclic stable games, in which evolution leads to an asymptotically stable set X^* where all types may co-exist even when they behave differently (again ignoring type 0). Now suppose that level-k types are randomly matched to play games that are drawn from a set of games $\mathcal{G}$. For the type space $K\backslash\{0\}$ it is possible to a general result for the case when $\mathcal{G}$ consists of one type-acyclic game and one type-cyclic game. If type 0 is included in the analysis general results are harder to come by. I prove a result for the case when $\mathcal{G}$ consists of (i) a Travelers' Dilemma, (ii) a 2×2-game with a unique interior ESS, and (iii) a 2×2 coordination game.

Before the desired results can be stated some more notation is needed: According to definition 2, in a type-cyclic game, k^c is the smallest number k such that $\beta^k(U) = \beta^{k'}(U)$ for some $k' < k$. Now let k^{c*} denote the smallest such k' so that $\beta^{k^{c*}}(U) = \beta^{k^c}(U)$. Thus type k^{c*} is the lowest type whose behavior is part of a cycle that is repeated for all higher types: Type $k^{c*} + i$ plays the same strategy as type $k^c + i$ for any $i \geq 0$. We also need to modify the definitions of $\tilde{K}$ and $\tilde{X}$. Let $\tilde{K}^{\mathcal{G}}$ be the set of types that behave like the highest type in all games in $\mathcal{G}$, $\tilde{K}^{\mathcal{G}} = \{k \in K : \beta^k(U) = \beta^\kappa(U), \forall G \in \mathcal{G}\}$, and let $\tilde{X}^{\mathcal{G}}$ be the set of states where only these types exist; $\tilde{X}^{\mathcal{G}} = \{x \in \Delta(K) : \sum_{k \in \tilde{K}^{\mathcal{G}}} x_k = 1\}$.

Proposition 5 ***(a)*** *Suppose that level-k types $K\backslash\{0\}$, play games from a set $\mathcal{G} = \{G^A, G^C\}$, where G^A is a type-acyclic game and G^C is a type-cyclic game, with associated numbers k^{NE} and k^c, probabilities μ^A and μ^C, and type game payoffs w^A and w^C, respectively.*

1. *If $\kappa - k^{NE} \geq k^{c*} - k^c$ then $\tilde{X}^{\mathcal{G}}$ is not asymptotically stable.*

2. *If* $\kappa - k^{NE} < k^{c*} - k^{c}$ *then* $\tilde{X}^{\mathcal{G}} = \{x \in \Delta(K) : x_{\kappa} = 1\}$ *is asymptotically stable if, for all* $k \in \{1, ..., k^{NE}\}$,

$$\mu^{A}\left(w^{A}(k,\kappa) - w^{A}(\kappa,\kappa)\right) < \mu^{C}\left(w^{C}(\kappa,\kappa) - w^{C}(k,\kappa)\right). \tag{8}$$

(b) *Suppose that level-k types* $K = \{0, 1, 2\}$ *play games from a set* $\mathcal{G}' = \{G^{TD}, G^{HD}, G^{CO}\}$, *consisting of a Travelers' Dilemma,* G^{TD}, *with parameters* $R = 3/2, P = 1/3$, $n \geq 4$, *a* 2×2 *stable game,* G^{HD}, *with matrix (6), and a* 2×2 *coordination game,* G^{CO}, *with matrix (5). If* $3\mu^{HD} < \mu^{TD}$ *then only the state where* $x_2 = 1$ *is asymptotically stable. If* $3\mu^{HD} > \mu^{TD}$ *then there is a unique asymptotically stable state, with the whole interior as basin of attraction, in which* $x_0 = 0$ *and*

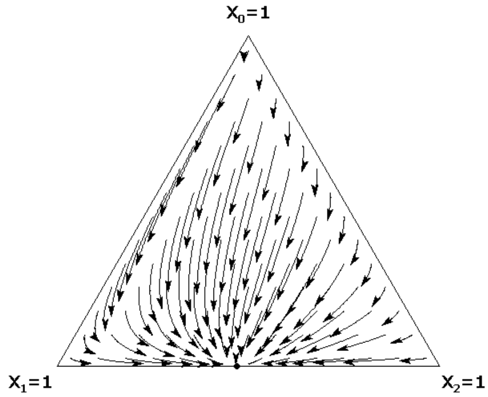
$$x_1 = \frac{6b\mu^{HD} - 2\mu^{TD}}{6\mu^{HD}(b+1) + \mu^{TD}}. \tag{9}$$

To see the logic behind part (a.1) of the above proposition, note that if κ is sufficiently large, then there are many types that play the same strategy as the highest type in the type-acyclic game, and some of these types play a best response to the highest type in the type-cyclic game. Thus in any state where everyone behaves like the highest type, there is some other mutant type that would earn strictly more than the incumbent types.

To see the intuition behind part (a.2) of the above proposition, consider the case of $\kappa = k^{NE} + k^{c}$, $k^{c*} = 1$ Under these assumptions the best response functions of the type games based on the type-acyclic game, and the type-cyclic game, only differ with respect to what type is the best reply to type κ. In both the type game based on the type-acyclic game and the type game based on the type-cyclic game type $k+1$ is the unique best response to type k for all $k \in \{1, 2, ..., \kappa - 1\}$. In the type game based on the type-acyclic game κ is the unique best response to itself, whereas in the type game based on the type-cyclic game, type 1 is the unique best response to type κ. The condition (8) assures that type κ is the unique best reply to itself also in the type game based on the combination of these games.

Figure 3 illustrates part (b) of the above proposition, for $K = \{0, 1, 2\}$ when each game is played with equal probability. The parameters of the Travelers' Dilemma and the stable game are the same as in figures 1 and 2, and in the coordination game $a = 1$. There is convergence to a unique state where types 1 and 2 co-exist.

Figure 3 *Evolution across games played by level-k types.* The set of games consist of three games, each played with the same probability: (i) Travelers' Dilemma with $R = 3/2$, $P = 1/3$, and $n = 6$. (ii) 2×2 stable game with $b = 1$. (iii) 2×2 coordination game with $a = 1$. The type space is $K = \{0, 1, 2\}$. The state $x = (0, 4/13, 9/13)$ (black dot) is asymptotically stable and has the whole interior as its basin of attraction.



4 Extensions

4.1 Heterogeneous Fictitious Play

So far the focus has been on the evolution of theories of mind that are used to predict opponent's initial behavior. This section studies the evolution of theories of mind used in the process of learning. People may then use their information about past play to predict future play. Fictitious play postulates that all individuals believe that the future will be like the past, and best respond to the average of past play. I modify this model and assume that there is a hierarchy of types $K = \{1, 2, ..., \kappa\}$, such type k plays a k times iterated best response to the average of past play. I will refer to this as the *heterogeneous fictitious play (HFP)* model. Note that there is no type 0 in this model.

Suppose that during her lifetime each individual is randomly matched to play a symmetric two-player normal form game G, τ times with τ different individuals from the same population. The average payoff over these τ interactions serves as fitness payoff in the evolutionary process. In order to keep things tractable I assume that all individuals of type 1 have a common prior with full support, and that higher types know about this prior. Let $h^t \in \Delta(S)$ be the aggregate play in period t. According to fictitious play the belief of type 1, γ^t, evolves as

$$\gamma^t = \frac{1}{t}\left(h^{t-1} + (t-1)\gamma^{t-1}\right), \ t \geq 2.$$

Note that $\{h^t\}_{t=1}^{\tau}$ and $\{\gamma^t\}_{t=1}^{\tau}$ are fully determined by γ^1 and x. Type 1 plays strategy $\sigma(1,\gamma^t) = \bar{\beta}(\gamma^t)$, and type k plays strategy $\sigma(k,\gamma^t) = \bar{\beta}^k(\gamma^t)$. The expected payoff of type k, against type k', in period t, is $\tilde{\pi}(\sigma(k,\gamma^t), \sigma(k',\gamma^t))$. Averaging over the τ periods, and recalling that $\{\gamma^t\}_{t=1}^{\tau}$ is fully determined by γ^1 and x, one gets

$$\Pi_k\left(k', x, \gamma^1\right) = \frac{1}{\tau}\sum_{t=1}^{\tau} \tilde{\pi}\left(\sigma\left(k,\gamma^t\right), \sigma\left(k',\gamma^t\right)\right).$$

The expected payoff of type k, in state x is

$$\Pi_k\left(x,\gamma^1\right) = \sum_{k'\in K} \Pi_k\left(k', x, \gamma^1\right) x_{k'}.$$

This is the evolutionarily relevant payoff. The payoffs will generally not be continuous in the state. As a consequence, the vector field (2) will generally not be Lipschitz continuous in the state. However, the vector field will be Lipschitz continuous almost everywhere. The reason is that behavior only changes in states where some type is indifferent between two or more strategies. This set is constituted by the union of a finite set of hyperplanes in the type space $\Delta(K)$. These hyperplanes divide the type space into a finite number of open set. Within each of these sets, behavior is constant across states. Therefore we can use the notion of a Filippov (1960) solution. (See Ito (1979) for a more accessible statement.)

Definition 8 *Consider the system $\dot{x} = \varphi(x)$ where φ is a real bounded measurable function, defined almost everywhere on a set $Q \subseteq \mathbb{R}^n$. For any x, let*

$$C(\varphi(x)) = \cap_{\delta:\delta>0} \cap_{Z:\mu(Z)=0} \overline{co}\left(\varphi\left(\bar{B}_\delta(x) \backslash Z\right)\right),$$

*where Z is an arbitrary set in $\mathbb{R}^n$, $\bar{B}_\delta(x)$ is the closed δ-ball around x, $\overline{co}$ denotes the closed convex hull, and μ is a Lebesgue measure. A **Filippov solution** to the system $\dot{x} = \varphi(x)$ with initial condition x^0, is an absolutely continuous function $\xi(\cdot, x^0) : T \to Q$ with $\xi(0,x^0) = x^0$, such that*

$$\frac{\partial}{\partial t}\left(\xi\left(t,x^0\right)\right) \in C\left(\varphi\left(\xi\left(t,x^0\right)\right)\right),$$

holds almost everywhere.

Intuitively, the set $C(\varphi(x))$ is a "cleaned up version" of $\varphi(x)$, constructed in the following way: From vectors associated with the ball $\bar{B}_\delta(x)$ we take away all those "strange" vectors that are only associated with some (Lebesgue) measure zero subset of the ball (any set Z such that $\mu(Z) = 0$). Then we take the convex hull of the remaining non-strange vectors. Finally we take the limit as we shrink the ball (intersection for all $\delta > 0$). The useful thing about the Filippov solution is that it does not have to respect the direction of the vector field on a measure zero set. Filippov showed that a solution in the above sense always exists. Our vector field satisfies the conditions of being real, bounded, and measurable. Furthermore it is defined everywhere on $\Delta(K)$. Hence our system has at least one Filippov solution. It turns out that this is all we need to prove the results we want.[16]

For a given game and a prior γ^1, let $\tilde{K}(\gamma^1)$ be the set of types that behave like the highest type in all periods;

$$\tilde{K}\left(\gamma^1\right) = \{k \in K : \bar{\beta}^k\left(\gamma^t\right) = \bar{\beta}^\kappa\left(\gamma^t\right), \forall t \leq \tau\}. \tag{10}$$

Let $\tilde{X}(\gamma^1)$ be the set of states where only these types exist – analogous to the definition of $\tilde{X}$ in equation (4). For WBRD-games have the following result:

Proposition 6 *Suppose that an underlying game satisfying WBRD is played by HFP types. For any κ, if $\beta(\gamma^1)$ is a singleton, then evolution from any interior initial condition converges to the unique asymptotically stable set $\tilde{X}(\gamma^1)$.*

For the case of 2×2 stable games one can prove result that is similar to those presented for the level-k model above in propositions 3 and 4. Here will investigate the following notorious example due to Shapley (1964);

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{11}$$

This game has a unique Nash equilibrium in which each of the three strategies are given equal weight. However, fictitious play does not converge to this equilibrium.

[16]Within evolutionary game theory differential inclusions are often analyzed with the help of Caratheodory solutions (e.g. Lahkar and Sandholm (2008)). A Caratheodory solution to $\dot{x} = \varphi(x)$ is an absolutely continuous function $\xi\left(\cdot, x^0\right) : T \to Q$ with $\xi\left(0, x^0\right) = x^0$, such that $\frac{\partial}{\partial t}\left(\xi\left(t, x^0\right)\right) = \varphi\left(\xi\left(t, x^0\right)\right)$, holds almost everywhere. For the present purposes the crucial difference between the concepts can be seen from the following example: Consider a vector field $\varphi : \mathbb{R}^2 \to \mathbb{R}$ such that $\varphi(x) = (1, 0)$ if $x_1 \neq 0$, and $\varphi(x) = (0, 1)$ if $x_1 = 0$. For an initial condition $x^0 = \left(0, x_2^0\right)$ there are two Caratheodory solutions, namely $\xi\left(t, x^0\right) = \left(t, x_2^0\right)$ and $\xi'\left(t, x^0\right) = \left(0, t - x_2^0\right)$. Only the former is a Filippov solution, since the set $\{x \in R^2 : x_1 = 0\}$ has measure zero. The reason that I prefer to use Filippov solutions in this paper is that they allow me to prove results that hold for all Filippov solutions, whereas I would generally not be able to prove that the same results hold for all Caratheodory solutions.

Instead play goes round in a cycle whose time average does not correspond to the Nash equilibrium. Intuitively we might think that if real humans were engaged in this game and initially behaved in accordance with fictitious play, they would eventually be able to detect the cycles and best respond to it, and thereby break out of the cycle. The present model of heterogeneous fictitious play is able to do justice to these intuitions. Type i behaves like type i modulo 3, abbreviated $i_{\text{mod}\,3}$.

Proposition 7 *Suppose that an underlying Shapley game is played by HFP types. For any $\kappa \geq 3$, if $\beta\left(\gamma^1\right)$ is a singleton then evolution from any interior initial state converges to the state where*

$$\sum_{i\in\{k\in K:k=1_{\text{mod}\,3}\}} x_i = \sum_{i\in\{k\in K:k=2_{\text{mod}\,3}\}} x_i = \sum_{i\in\{k\in K:k=3_{\text{mod}\,3}\}} x_i = 1/3,$$

and aggregate behavior corresponds to the unique Nash equilibrium in all periods.

4.2 Partially Observed Level-k Types

As explained in the introduction, there are many situations in which people play an unfamiliar game with people they already know something about. In particular players may have information about their opponents' theories of mind. I propose a simple way of extending the level-k model to the case of partially observed types. The only modification that I add to the level-k model is that when an individual of type k faces an opponent of a lower type $k' < k$ then the former is able to understand how the latter thinks, and hence former best responds to the latter. The lower type is assumed to behave exactly as in the ordinary level-k. Formally, type $k \in K$ plays

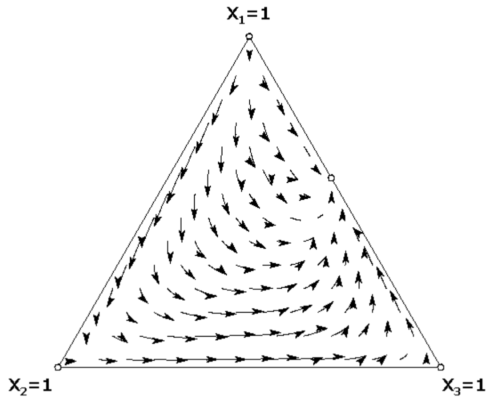
$$\sigma\left(k,k'\right) = \begin{cases} \beta^{k'+1}\left(U\right) & \text{against } k' \in \{0,1,2,...k-1\} \\ \beta^{k}\left(U\right) & \text{against } k' \geq k \end{cases}.$$

Since the behavior of a type depends on what type she encounters, equation (1) is replaced by the following expression for the expected payoff of type k, in state x,

$$\Pi_k^G\left(x\right) = \sum_{k'\in K} x_{k'}\tilde{\pi}\left(\sigma\left(k,k'\right),\sigma\left(k',k\right)\right).$$

Of course, this is very simplified account of what happens when one individual understands how another individual reasons while the latter does not understand how the former reasons. Still, this specification lends itself to straightforward analysis. Introducing partial observability into the level-k model changes the results sub-

Figure 4 *Travelers' Dilemma played by partially observed level-k types.* Parameters $R = 3/2$, $P = 1/3$, and $n = 6$. Type space $K = \{1, 2, 3\}$. Evolution from any interior initial state converges to the state $x = (4/7, 0, 3/7)$ (white dot) which is not stable.



stantially compared with the standard level-k model: The set $\tilde{X}$ may be unstable even if the underlying game is type-acyclic and $\kappa \geq k^{NE}$. If fact, this might happen even if the game satisfies WBRD, as demonstrated by the Travelers' Dilemma:

Proposition 8 *Suppose that an underlying Travelers' Dilemma game, is played by partially observed level-k types* $K \backslash \{0\}$. *For any* κ, *if* $P \in (p-1, p)$ *for some* $p \in \mathbb{N}$, *then every state where* $x_k = x_{k-1} = ... = x_{k-p} = 0$ *for some* $k > p$, *is unstable.*

The claim in the above proposition implies that the set of state where everyone behaves like the highest type is unstable. The intuition for this result is that lower strategies are more destructive, so that when higher types meet each other, they earn less than what lower types earn when they meet higher types. A lower type is committed to a less destructive strategy, and may thereby induce a higher type to choose a less destructive strategy, something that might benefit both types. When there is a large fraction of the high type, this mechanism favors the growth of the low type. Figure 4 illustrates what happens in the case of $K = \{1, 2, 3\}$, with the same parameters as in previous figures. Evolution from any interior initial state converges to the state $x = (4/7, 0, 3/7)$, which is not stable. Adding a type 0 would not alter this result since type 0 would become extinct asymptotically.

The next proposition spells out the diametrically different effects that introducing partial observability may have in two different kinds of cyclic games. We say that a game with $n \geq 3$ is *monocyclic* (Hofbauer (1995)) if $s' \neq s''$ and $s' \neq \beta(s'')$ implies $\pi(s, s) > \pi(s', s'')$, still assuming that β is single-valued. Monocylic games have the special property that each pure strategy is the second best response to itself.

Proposition 9 *Suppose that an underlying game is played by partially observed level-k types.* ***(a)*** *If the types space is* $K \backslash \{0\}$*, for any* κ*, and if the underlying game is monocyclic, then evolution from any interior initial state converges to the unique asymptotically stable state where* $x_\kappa = 1$*.* ***(b)*** *If the type space is* K*, for any* $\kappa \geq 3$*, and if the underlying game is a* 2×2*-game with an interior ESS, with payoffs as in matrix (6), then evolution from any interior initial state converges to a unique asymptotically stable state where* $x_0 = 0$*, and* $x_i = bx_j$*, for any odd number* $i \leq \kappa$*, and any even number* $j \leq \kappa$*.*

Part (a) says that for some specific completely cyclic games introducing partial observability creates a strict advantage for higher types: Regardless of how large κ is, evolution will lead to the state where only type κ exists. Thus if $\kappa \to \infty$ we get evolution towards *infinitely high types*, something that was not possible with unobserved types. Part (b) states a result that goes in the opposite direction. For any κ, evolution leads to a state where all types, except type 0, co-exist. As $\kappa \to \infty$ we get evolution towards *infinite diversity*, except for the extinction of type 0.[17]

5 Discussion

5.1 Evolution of Initial Responses

There are only a few papers studying the evolution of cognitive types. A pioneering paper is Stahl (1993). In his model there is a set of types $n \in \{0, 1, 2, ...\}$. Type 0 is divided into subtypes, each programmed to a different pure strategy. Type n believes that everyone else is of a lower type, is able to deduce what lower types will do, and chooses among strategies that are n^{th} order rationalizable conditional on the actual distribution of types. In order to choose among the strategies that are n^{th} order rationalizable, each individual has a secondary strict preference ordering over strategies. He finds that programmed individuals may survive since "being right is just as good as being smart". Banerjee and Weibull (1995) study the interaction between individuals that are preprogrammed to different strategies and individuals that optimize given a correct belief about the strategy of the opponent (full information case) or the population distribution of strategies (incomplete information case). Another related paper is Stennek (2000) who studies the evolutionary advantage of ascribing different degrees of rationality to one's opponent. An individual of type $d \in \{0, 1, 2, ...\}$ believes that everyone else is of type $d-1$ and

[17] The proof exploits the fact that the off-diagonal elements are zero. Hence the result does not extend to all games that are strategically equivalent to matrix (6).

chooses some d-iterations undominated action, in accordance with some preference over the pure strategies. There are at least three important differences between the models in these papers and the present model: First, all of these papers assume that a fixed game is played recurrently and that some individuals are programmed to pure strategies. The fact that some individuals are programmed to pure strategies makes it difficult to study evolution of initial responses and evolution across games, since it is not clear how such individuals transfer their behavior between games. Moreover, it would seem that in order to survive such programmed individuals have to pick the right strategies in all the games they play. Secondly, these papers build on behavioral models that lack the kind of empirical support that the level-k and cognitive hierarchy models have (see Costa-Gomes and Crawford (2006) and Camerer (2003)). Third, all of these models include many types whose behavior is not fully determined by best response given beliefs. Such an approach potentially confounds the question of how theories of mind have evolved, and the question of how optimizing behavior has evolved.

Samuelson (2001a) studies the evolution of finite automata used to implement strategies in a class of three different games; an ultimatum game, an infinite horizon bargaining game, and a tournament. Each individual uses a single automaton to implement play in all three games. There is a cost that is increasing in the number of states of the automata. Thus there is an incentive to save on states and thereby behave in a way that is not tailored to each of the games – e.g. to behave in the same way in the ultimatum game as in the infinite horizon bargaining game. In a similar way, in the present paper I assume that an individual's type is constant across games. In reality an individual's type is likely to vary somewhat over time and across games, but as long as each individual's type only varies within some limits, one may expect the results presented in this paper to be of relevance.

Sgroi and Zizzo (2009) study how a player endowed with a neural network adapts her play across different 3×3-games, assuming that all other players play a Nash equilibrium strategy. Their simulation experiments are broadly consistent with a level-k model.

5.2 Evolution of Modes of Learning

As mentioned in the introduction there is almost no literature on evolution of different learning rules. Heller (2004) studies competition between, on the one hand individuals that are preprogrammed to pure strategies, and on the other hand, individuals who play a best response to the opponent she faces. The latter behavior is intended to be a reduced form for how a learner would behave after learning has taken place. Since the environment changes stochastically the learners may survive even though they incur a strictly positive cost. Josephson (2008) uses simulations to compare fictitious play and reinforcement learning. He finds that evolution may end up putting roughly equal weights on these two modes of learning. Both studies are restricted to sets of games with identical strategy spaces.

Stahl's (2000) rule learning model assumes that each individual is endowed with propensities for different learning rules. One rule is to imitate past behavior, another one is to play a best response to past play. Further rules specify some iterated best response to past play. The propensities to follow these different rules are updated in relation to how well they perform in the game under study. This is undoubtedly an interesting model for explaining and predicting how people learn to play experimental games, but Stahl does not provide any general results on which propensities that may be evolutionarily stable in different classes of games. The current paper can be viewed as a step in that direction.

5.3 The Social Brain and Cognition Costs

According to the prominent "social brain", or "Machiavellian intelligence", hypothesis, the extraordinary cognitive abilities of humans evolved as a result of the demands of social interactions, rather than the demands of the natural environment (Humphrey (1976), for an introduction see Dunbar (1998)). In a single person decision problem there is a fixed benefit of being smart, but in a strategic situation it may be important to be smarter than the opponent. Robson (2003) models the Machiavellian intelligence hypothesis as the interaction between an uninformed player and an informed player. The informed player does not want to reveal her information, but the uninformed player wants her to do so. Both players use noisy bounded recall strategies. It turns out that for any equilibrium, each player would benefit from getting a longer recall. Thus there is a always pressure towards strategic sophistication in the form of greater recall. The results in this paper complement the social brain hypothesis, and Robson's results, by suggesting mechanisms that may sustain heterogeneity with respect to theory of mind abilities – in a way that is consistent with experimental findings.

From an evolutionary perspective, the potential advantage of a better theory of mind has to be traded off against the cost of increased reasoning capacity. Increased cognitive sophistication, in the form of higher order beliefs, is probably associated with non-negligible costs (Holloway (1996), Kinderman et al. (1998)). Such costs have been excluded from the formal analysis. However, the potential effect of cognitive costs should be kept in mind when interpreting the findings. Adding cognitive costs will increase the possibilities for lower types to survive. There does not seem to be any reasons to expect that adding costs would make the evolution of a heterogeneous population less likely. As a brief illustration of this point, suppose that the type game without cognitive costs can be represented by the negative of the identity matrix. In this case the type game without cognitive costs is strictly stable, with an interior ESS, i.e. $v \cdot (-\mathbf{I})\, v < 0$ for all $v \in \mathbb{R}_0^{\kappa+1} = \{v \in \mathbb{R}^{\kappa+1} : \sum v_i = 0\}$, $v \neq \mathbf{0}$. Hence evolution leads to a state where all types co-exist. This conclusion will not be altered if we add a cognitive cost c_k for each type k. To see why, let $\mathbf{c} \in \mathbb{R}^{\kappa+1}$ be the vector of costs and note that $(\mathbf{c1}')\, v = \mathbf{0}$ for all $v \in \mathbb{R}_0^{\kappa+1}$ so that $v \cdot (-\mathbf{I} - \mathbf{c1}')\, v < 0$. Moreover, if some types become extinct due to the added costs the remaining types will still play a type game whose payoff matrix can be written as $-\hat{\mathbf{I}} - \hat{\mathbf{c}}\hat{\mathbf{1}}'$, for an identity matrix $\hat{\mathbf{I}}$, and vectors $\hat{\mathbf{c}}$ and $\hat{\mathbf{1}}$, of a dimensionality that corresponds the number of surviving types.

The results presented in this paper are qualitative in nature. They indicate mechanisms that may lead to the evolution of heterogeneous and bounded depth of reasoning. However, they cannot be used directly to make quantitative predictions about the distribution of types. Empirically, most experimental subjects seem to behave as if they are of type 1 or 2, and individuals of type 3 and above are rare (Costa-Gomes and Crawford (2006), Camerer (2003)). This distribution is probably the outcome both of the mechanisms investigated in this paper, and of costs of cognition.

In this paper, preferences have been taken as given and focus has been on how evolution shapes cognition given preferences. There is a large literature on preference evolution, which takes rationality and cognition as given. An interesting possible avenue for future research is to combine these approaches and study the co-evolution of preferences and cognition.

6 Conclusion

This paper has undertaken an evolutionary analysis both of the level-k model of initial play, and of a heterogeneous fictitious play model of learning. The analysis includes evolution of types across games. Furthermore the paper has extended the level-k model to the case of partially observed types. It was found that an evolutionary process, based on payoffs earned in different games, both with and without partial observability, could lead to a polymorphic population where relatively unsophisticated types survive.

7 Appendix: Proofs

7.1 The Level-k Model

Proof of Proposition 1. Suppose that the underlying game is type-acyclic with $\kappa \geq k^{NE}$. Assume without loss of generality $k^{NE} = \kappa$, so that $\tilde{X}$ only contains the state where $x_\kappa = 1$. It follows that (κ, κ) is a pure Nash equilibrium of the type game. By the definition of k^{NE} we have $\beta^k(U) \neq \beta^{k+1}(U)$ for all $k < k^{NE} = \kappa$. Hence the profile (κ, κ) is a strict equilibrium, implying that κ is an ESS in the type game. Thus the state where $x_\kappa = 1$ is asymptotically stable.

Suppose that the underlying game is type-cyclic with $\kappa \geq k^c$. We have $\beta(\beta^\kappa(U)) = \beta^k(U)$ for some $k < \kappa$, and this best reply is strict, so the set $\tilde{X} = \{x \in \Delta(K) : x_\kappa = 1\}$ is not stable. ■

7.1.1 Type-Acyclic Games

Lemma 1 *If a game satisfies WBRD then it is type-acyclic.*

Proof. In a completely cyclic game there is some s and some k such that $\beta^k(s) = s$, and $\beta^i(s) \neq \beta^j(s)$ for all $i, j \in \{0, ..., k-1\}$ with $j \neq i$. This implies that $\pi(\beta^i(s), \beta^i(s)) < \pi(\beta^{i+1}(s), \beta^i(s))$ for all $i \in \{0, ..., k-1\}$ and $\pi\left(\beta^{k-1}(s), \beta^{k-1}(s)\right) < \pi\left(s, \beta^{k-1}(s)\right)$. In contrast, in a WBRD-game we have $\pi\left(s, \beta^{k-1}(s)\right) \leq \pi\left(\beta(s), \beta^{k-1}(s)\right) \leq ... \leq \pi\left(\beta^{k-2}(s), \beta^{k-1}(s)\right) \leq \pi\left(\beta^{k-1}(s), \beta^{k-1}(s)\right)$. Thus a WBRD-game is not type-cyclic, so it is type-acyclic. ■

Proof of Proposition 2. (i) *Payoffs*: Since all types $k \geq k^{NE}$ behave like type κ, assume $\kappa = k^{NE}$. Generalizing to $\kappa \geq k^{NE}$ is straightforward. The payoff matrix of the type game with types $K\backslash\{0\}$ is

$$\mathbf{A} = \begin{pmatrix} w(1,1) & w(1,2) & . & w(1,\kappa) \\ w(2,1) & w(2,2) & . & w(2,\kappa) \\ . & . & . & . \\ w(\kappa,1) & w(\kappa,2) & . & w(\kappa,\kappa) \end{pmatrix}.$$

Since $\beta^{\kappa-1}(U) \neq \beta^\kappa(U)$ we have $\beta^k(U) \neq \beta^{k-1}(U)$ for all $k \leq \kappa$, which implies $w(k+1, k) > w(k, k)$ for all $k \leq \kappa - 1$. By WBRD we have $w(k+1, i) \geq w(k, i)$ for all $i \geq k$.

(ii) *Convergence*: The type game can be solved by iterated elimination of weakly dominated strategies. More precisely, type $k+1$ weakly dominates type k once all types $i < k$ have been eliminated. Furthermore, it holds that $w(k+1, k) > w(k, k)$.

Mohlin (2011) shows that under these conditions the replicator dynamic converges to the weak dominance solution. For the convenience of the reader a specialized version of this result is provided in the supplement as well (lemma S3).

(iii) *Stability*: Follows from the fact that the game is type-acyclic. ■

For the analysis of the Travelers' Dilemma we need the following lemma.

Lemma 2 *In the Travelers' Dilemma, let $R+P=l+r$ for some $l \in \mathbb{N}$ and some $r \in (0,1)$. The best response against the uniform distribution is $\beta(U) = n-l-1$.*

Proof. This is only a matter of calculations. Available upon request. ■

Proof of Corollary 1. Trivial, therefore omitted. ■

7.1.2 Type-Cyclic Games

Lemma 3 *Let $G=(2,S,\pi)$ be a symmetric normal form 2-player game. If G is stable then the 2-player game $G'=(2,S',\pi)$ that is generated by restricting the strategy set to $S' \subseteq S$, is stable.*

Proof. Without loss of generality assume that the strategies of G are named $\{1,2,..,n\}$ and let $\mathbf{A}$ be the payoff matrix associated with game G. If $v \cdot \mathbf{A}v < 0$ for all $v \in \mathbb{R}_0^n$, $v \neq \mathbf{0}$, then $v \cdot \mathbf{A}v < 0$ for all $v \in \mathbb{R}_0^{n'} = \{v \in \mathbb{R}^n : \sum_i v_i = 0$ and $v_i = 0$ iff $i \notin S'\} \subseteq \mathbb{R}_0^n$, $v \neq \mathbf{0}$. Thus G' is stable. ■

Proof of Proposition 3. First, suppose that no two types play the same strategy. Since the underlying game is stable the type game is also stable by lemma 3, and so it has a unique ESS. In a stable game the state where aggregate behavior corresponds to the unique ESS is asymptotically stable under the replicator dynamic, and has the whole interior as its basing of attraction (see e.g. Sandholm (2011), chapter 6). Thus evolution in the type game (where all types play different strategies) leads to a unique asymptotically stable state where aggregate behavior corresponds to the ESS of the type game. It is easy to see that if some types play the same strategy then we have a unique asymptotically stable *set* instead of a unique asymptotically stable *state*.

1. Since the underlying game is stable with an interior ESS it is type-cyclic. To obtain a contradiction suppose that there is some $x \in X^* \cap \tilde{X}$. Since $x \in \tilde{X}$ every type k with $x_k > 0$ plays the same strategy s. Since $\kappa \geq k^c$ there exists some other type k', with $x_{k'} = 0$, that plays $\beta(s)$. Hence a mutant of type k' earns strictly more than the incumbent types.

We prove points 2 and 3 in reverse order:

3. If the game is completely cyclic and $\kappa = n$ then each strategy in the underlying game is played by exactly one type so that the unique ESS of the type game is identical to the ESS of the underlying game.

2. If the game is completely cyclic and $\kappa \geq n$ then some types play the same strategy. Thus instead of a unique state where aggregate behavior corresponds to the ESS of the underlying game we have a set of states where aggregate behavior corresponds to the ESS of the underlying game. ■

Lemma 4 *If G is a potential game then the type game generated by types K playing G is potential.*

Proof. Since the underlying game is potential the payoff to strategy i against strategy j can be decomposed as $a_{ij} = c_{ij} + r_j$. We need to show that the payoffs of the type game can be similarly decomposed: The payoff to type $k \geq 1$, playing strategy i, against type $k' \geq 1$, playing strategy j, can be written as $w(k, k') = a_{ij} = c_{ij} + r_j$, where $c_{ij} = c_{ji}$ for all i, j. The payoff to type $k \geq 1$, playing strategy i, against type 0, can be written as

$$w(k, 0) = \frac{1}{n}\sum_{j=1}^{n} a_{ij} = \frac{1}{n}\sum_{j=1}^{n} c_{ij} + \frac{1}{n}\sum_{j=1}^{n} r_j = c_{i0} + r_0.$$

The payoff to type 0, against type $k' \geq 1$, playing strategy j, can be written as

$$w(0, k') = \frac{1}{n}\sum_{i=1}^{n} a_{ij} = \frac{1}{n}\sum_{i=1}^{n} c_{ij} + \frac{1}{n}\sum_{i=1}^{n} r_j = c_{0j} + r_j.$$

Finally the payoff of type 0 playing against type 0, can be written as

$$w(0, 0) = \frac{1}{n}\sum_{i}^{n}\sum_{j=1}^{n} a_{ij} = \frac{1}{n}\sum_{i}^{n}\sum_{j=1}^{n} (c_{ij} + r_j) = c_{00} + r_0.$$

■

Proof of Proposition 4. First, suppose that no two types $k, k' \geq 1$ play the same strategy. By lemma 4 the type game is potential. In a potential game the replicator converges to some asymptotically stable set from any interior initial condition (see e.g. Sandholm (2011), chapter 6). Thus evolution in the type game (where all types $k \geq 1$ play different strategies) leads to an asymptotically stable state. Again, if some types play the same strategy then we have convergence to a an asymptotically stable set instead of state.

1. The same argument as in the proof of proposition 3.1 establishes that if $\kappa \geq k^c$ then $X^* \cap \tilde{X} = \emptyset$.

2. Let H be the first strategy and D the second strategy in matrix 6. Note the following property of 2×2-games with a unique interior ESS: If $\sigma_s(x) > \sigma_s^{ESS}$ then strategy $s \in \{H, D\}$ earns more than strategy $s' \neq s$. Let x_H, x_D, and x_U denote the fractions of the population that plays H, D, and U, respectively. In all interior state we have $x_H > 0$, $x_D > 0$, and $x_U > 0$. It is trivial to see that no monomorphic states are stable. Suppose the system is initially in an interior state where $x_H + x_U/2 > \sigma_H^{NE}$. Then $\dot{x}_H < 0$ and $\dot{x}_D > \dot{x}_U > \dot{x}_H$, so x_H decreases, x_D increases, and x_U may increase or decrease. This process continues until asymptotically $x_H + x_U = \sigma_H^{NE}$. Similar reasoning applies if the system is initially in an interior state where $x_D + x_U/2 > \sigma_D^{NE}$. Thus evolution from any interior initial state converges to some state where $x_s + x_U/2 = \sigma_s^{NE}$. ■

7.1.3 Evolution across Games

Proof of Proposition 5. (a) We study the type game derived from the combination of G^A and G^C with probabilities μ^A and $\mu^C = 1 - \mu^A$.

1. Suppose that $\kappa - k^{NE} \geq k^c - k^{c*}$. The set $\tilde{K}^{\mathcal{G}}$ consists of the types $k \geq k^{NE}$ that play the strategy $\beta^{\kappa}(U)$ in the type-cyclic game. All types $k \geq k^{NE}$ earn the same in the type-acyclic game. Since $\kappa - k^{NE} \geq k^c - k^{c*}$ there is a type $k' = \kappa - (k^c - k^{c*}) \geq k^{NE}$ that plays a best reply to type κ, i.e. $\beta^{k'}(U) = \beta(\beta^{\kappa}(U))$ in the type-cyclic game. Since $\beta(\beta^{\kappa}(U)) \neq \beta^{\kappa}(U)$ in a cyclic game, we have $k' \notin \tilde{K}^{\mathcal{G}}$. Thus in any state $x \in \tilde{X}^{\mathcal{G}}$, a mutant of type $k' \notin \tilde{K}^{\mathcal{G}}$ earns more than each type $k \in \tilde{K}^{\mathcal{G}}$. Hence $\tilde{X}^{\mathcal{G}}$ is not asymptotically stable.

2. Suppose that $\kappa - k^{NE} < k^c - k^{c*}$. It follows that there is no type $k' \geq k^{NE}$ that play a best response to type κ in the type-cyclic game. Hence $\tilde{K}^{\mathcal{G}} = \{\kappa\}$. Condition (8) says that κ is a strict best reply to any $k < \kappa$ in the type game based on the combination of G^A and G^C. Hence $\tilde{X}^{\mathcal{G}} = \{x \in \Delta(K) : x_{\kappa} = 1\}$ is asymptotically stable.

(b) In the coordination game type 1 and 2 play $\beta(U)$, so the payoff matrix is

$$\begin{pmatrix} (a+1)/4 & a/2 & a/2 \\ a/2 & a & a \\ a/2 & a & a \end{pmatrix}, \tag{12}$$

so type 0 is strictly dominated. In the Hawk Dove game type 1 plays $\beta(U) = D$, and type 2 plays $\beta(D) = H$. Thus the payoff matrix is

$$\begin{pmatrix} -(1+b)/4 & -1/2 & -b/2 \\ -1/2 & -1 & 0 \\ -b/2 & 0 & -b \end{pmatrix}, \tag{13}$$

so type 0 earns the same as a mixed strategy in the type game, which puts equal probability on type 1 and 2. In the Travelers' Dilemma type 1 plays $n-l-1$, and type 2 plays $n-l-2$. Using $P = 1/3$ and $R = 3/2$ yields the payoff matrix

$$\begin{pmatrix} \tilde{\pi}(U,U) & \tilde{\pi}(U, n-l-1) & \tilde{\pi}(U, n-l-2) \\ \tilde{\pi}(n-l-1, U) & n-2 & n-\frac{10}{3} \\ \tilde{\pi}(n-l-2, U) & n-\frac{3}{2} & n-3 \end{pmatrix}. \tag{14}$$

One can verify that type 0 is dominated for all $n \geq 4$ (calculations available upon request).

The conclusion from these three games is that type 0 will be strictly dominated in the type game based on the combination of the games in $\mathcal{G}$, provided that $\mu^{HD} \neq 1$. Thus type 0 will be extinct so we disregard type 0 for the rest of the analysis. We can also disregard the payoffs from the coordination game (matrix 12) since type 1 and 2 earn the same in that game. It follows that we can restrict attention to the type game between type 1 and 2, derived from the Hawk Dove game and the Travelers' Dilemma. Deleting the payoffs involving type 0, and adding (13) and (14), weighted by μ^{HD} and μ^{TD} respectively, yields

$$\begin{pmatrix} w(1,1) & w(1,2) \\ w(2,1) & w(2,2) \end{pmatrix} = \begin{pmatrix} -\mu^{HD} - 2\mu^{TD} & -\frac{10}{3}\mu^{TD} \\ -\frac{3}{2}\mu^{TD} & -b\mu^{HD} - 3\mu^{TD} \end{pmatrix}. \tag{15}$$

We have $w(1,1) < w(2,1)$ for all μ. We have $w(1,2) > w(2,2)$ if and only if $3b\mu^{HD} > \mu^{TD}$. In that case the game (15) has a unique interior ESS where (9) holds. If $3b\mu^{HD} < \mu^{TD}$ then only $x_2 = 1$ is ESS, and hence the corresponding population state is asymptotically stable. ■

7.2 Heterogeneous Fictitious Play

Behavior depends on the state and on initial beliefs, so let $w(k, k', x, \gamma^1) = \Pi_k(k', x, \gamma^1)$. The proofs from above regarding the level-k model can be applied more or less directly to the *HFP* model, by using $w(k, k', x, \gamma^1)$ instead of $w(k, k')$ and showing that the payoff relations that obtain in the level-k model, obtain for the *HFP* model, for almost every state x. The reason that we can ignore measure zero sets of states is that we now study Filippov solutions.

Lemma 5 *Suppose that $\beta(s)$ is a singleton for all s and that $\beta(U)$ is a singleton. If γ^1 has full support and if $\beta(\gamma^1)$ is a singleton, then the set of states x that, together with γ^1, induce histories such that some type is indifferent between two or more strategies in some period, has measure zero.*

Proof of Lemma 5. The assumptions that $\beta(s)$ is a singleton for all s, and that $\beta(\gamma^1)$ is a singleton, implies that all higher types always will play pure strategies. Thus it is sufficient to show that the set of states that induce histories in which type 1 is indifferent, has measure zero.

(i) Since γ^1 has full support, so has γ^t for all t. First we show that the set of beliefs with full support, at which type 1 indifferent between two or more strategies, has measure zero: The set of beliefs γ at which type 1 is indifferent between two or more strategies is

$$\Gamma^I = \{\gamma \in \Delta(S) : \gamma \text{ has full support and } \exists s, s^* \in S \text{ s.t. } \tilde{\pi}(s,\gamma) = \tilde{\pi}(s^*,\gamma)\}.$$

The assumption that $\beta(U)$ is a singleton implies that for any pair of strategies s and s^* there is some s' such that $\tilde{\pi}(s,s') \neq \tilde{\pi}(s^*,s')$. This implies that the dimension of Γ^I is lower than the dimension of $\Delta(S)$. Thus Γ^I is a hyperplane in $\Delta(S)$, and as such it has measure zero.

(ii) Now we show that the set of states that induce histories where type 1 is indifferent, has measure zero. Recall $\gamma^t = (h_{t-1} + (t-1)\gamma^{t-1})/t$. Since Γ^I is a hyperplane it follows that if $\gamma^{t-1} \notin \Gamma^I$ then there is a measure zero set of states that induce aggregate behavior h_{t-1} such that $\gamma^t \in \Gamma^I$. Since we have assumed that $\beta(\gamma^1)$ is a singleton, an inductive argument establishes that, for each period, the set of states that induce indifference, given γ^1, has measure zero. Since the number of periods is finite it follows that the set of states that induce indifference in some period, given γ^1, has measure zero. ■

Proof of Proposition 6. By lemma 5 we can assume that all types play pure strategies in all states, without loss of generality. Let $s(k,\gamma^t)$ denote the pure strategy that is chosen by type k given the belief γ^t. Assume $\beta^{\kappa-1}(\gamma^1) \neq \beta^{\kappa}(\gamma^1)$ so that all types distinguish themselves behaviorally, at least in the first period. Extension to the case when several types behave like type κ already in the first period, is straightforward. Based on these assumption one can verify that the following payoff relations hold, analogous to what was obtained for the level-k model above (part (i) of the proof of proposition 2),

$$w\left(k+1,k,x,\gamma^1\right) > w\left(k,k,x,\gamma^1\right),$$
$$w\left(k+1,i,x,\gamma^1\right) \geq w\left(k,i,x,\gamma^1\right), \forall i \geq k.$$

The rest of the proof is identical to the proof of proposition 2. ■

Proof of Proposition 7. By lemma 5 we can assume that all types play pure strategies in all states. In period t type 1 plays $\beta\left(\gamma^t\right)$ and type k plays $\beta\left(\gamma^t\right) + (k+1)_{\text{mod}\,3}$. Thus in each period the payoffs to types $1_{\text{mod}\,3}$, $2_{\text{mod}\,3}$ and $3_{\text{mod}\,3}$ are given by the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \tag{16}$$

where the entry in the i^{th} row and the j^{th} column represents the payoff to type $i_{\text{mod}\,3}$ against type $j_{\text{mod}\,3}$. Since these payoffs are earned in each period, the above matrix (16) is the payoff matrix of the type game. Note that this matrix is the transpose of the payoff matrix (11) of the underlying game. It is straightforward to show that these matrices are negative definite with respect to the tangent space. Since the unique Nash equilibrium is interior it follows the game is strictly stable with an interior ESS. Thus the replicator dynamics converges from any interior initial state to the unique state $(1/3, 1/3, 1/3)$. ■

7.3 The Level-k Model with Partially Observed Types

Lemma 6 *Consider the Travelers' Dilemma, played by partially observed level-k types. If* $k \geq 1$ *then*

$$w(1,k) > w(2,k) > ... > w(k-1,k) \tag{a}$$
$$w(k-1,k) = w(k,k) - P \tag{b}$$
$$w(k,k) < w(k+1,k) \tag{c}$$
$$w(k+1,k) = ... = w(\kappa,k). \tag{d}$$

Proof. A matter of straighforward calculations. Available upon request. ■

Proof of Proposition 8. I prove the result for the case of $P \in (0,1)$. Generalization is straightforward. Consider a state x where $x_i = x_{i-1} = 0$ for at least one $i > 0$. There are three cases to consider: Either (I) there is at least one type $k < \kappa - 1$ such that $x_k = x_{k-1} = 0$ and $x_{k+1} > 0$, or (II) $x_{\kappa-1} = x_{\kappa-2} = 0$ and $x_\kappa > 0$, or (III) there is some $k \leq \kappa - 1$ such that $x_{k'} = 0$ for all $k' \geq k$, and $x_{k-1} > 0$.

Case I: Suppose that there is at least one type $k < \kappa-1$ such that $x_k = x_{k-1} = 0$ and $x_{k+1} > 0$. Using $x_k = 0$, we have

$$\begin{aligned}\Pi_{k-1}(x) - \Pi_{k+1}(x) = &\sum_{i=1}^{k-2} (w(k-1,i) - w(k+1,i))\, x_i \\ &+ (w(k-1,k+1) - w(k+1,k+1))\, x_{k+1} \\ &+ \sum_{i=k+2}^{\kappa} (w(k-1,i) - w(k+1,i))\, x_i.\end{aligned}$$

By lemma 6(d) the first term on the right hand side is zero. By lemma 6(b) the second term equals $(1-P)\, x_{k+1}$ so by the assumption that $P < 1$ this is positive. By lemma 6(a) the third term is strictly positive, so $\Pi_{k-1}(x) > \Pi_{k+1}(x)$. Thus a mutant of type $k - 1 < \kappa - 2$ entering the population will earn more than type $k + 1 < \kappa$.

Case II: Suppose $x_{\kappa-1} = x_{\kappa-2} = 0$ and $x_\kappa > 0$. By logic similar to that in case I we have

$$\Pi_{\kappa-2}(x) - \Pi_\kappa(x) = \sum_{i=1}^{\kappa-3} (w(\kappa-2,i) - w(\kappa,i))\, x_i + (w(\kappa-2,\kappa) - w(\kappa,\kappa))\, x_\kappa,$$

where the first term on the right hand side is equal to zero, by lemma 6(d) and, the second term is at least $(1-P)\, x_\kappa > 0$, by part (a) and (b) of lemma 6. Thus a mutant of type $\kappa - 2$ will earn more than type κ.

Case III: Suppose that there is some $k \leq \kappa - 1$ such that $x_{k'} = 0$ for all $k' \geq k$, and $x_{k-1} > 0$. The difference in average payoff to types k and $k - 1$ is

$$\Pi_k(x) - \Pi_{k-1}(x) = \sum_{i=1}^{k-2} (w(k,i) - w(k-1,i))\, x_i + (w(k,k-1) - w(k-1,k-1))\, x_{k-1}.$$

By lemma 6(d), the first term on the right hand side is zero and by lemma 6(c) the second term is strictly positive, so $\Pi_k(x) > \Pi_{k-1}(x)$. Thus a mutant of type k entering the population will earn more than type $k - 1$. ■

Proof of Proposition 9. **(a)** Since higher types play a unique best response to lower types, it follows that in each column of the payoff matrix of the type game, all entries below the diagonal are the same, and they are larger than the diagonal entry. By monocylicity, all entries above the diagonal in the type game are lower than the diagonal entries. Thus higher types strictly dominate lower types, and evolution from an interior initial condition converges to the state with $x_\kappa = 1$.

(b) **(i)** *Behavior and payoffs*: Let the first strategy be H and the second strategy be D. We have $\beta(U) = D$. Type $k \geq 1$ plays H against an odd $k' < k$ and D against an odd $k' < k$. An odd k plays D against $k' \geq k$ and an even k plays H against $k' \geq k$. This results in the following type game

$$\begin{pmatrix} -(1+b)/4 & -1/2 & -1/2 & . \\ -1/2 & -1 & 0 & . \\ -1/2 & 0 & -b & . \\ . & . & . & . \end{pmatrix}. \tag{17}$$

(ii) *Extinction of type 0*: Let $z = \{z_i\}_{i=1}^{\kappa}$ denote a mixed strategy in the type game represented by (17). Such a strategy z dominates type 0 in the type game if and only if $z_i < 1/2$ for all odd i and $z_i < 1/2b$ for all even i. Thus, if κ is odd then there is a strategy z that dominates type 0 if and only if

$$\frac{1}{2}\frac{\kappa+1}{2} + \frac{1}{2b}\frac{\kappa-1}{2} > 1,$$

or equivalently $\kappa > (3b+1)/(b+1)$. Note that the right hand side is increasing in b and approaches 3 as $b \to \infty$. Thus, for any finite b it holds that $(3b+1)/(b+1) < 3$. Thus the corresponding strategy in the type game will asymptotically become extinct under the replicator dynamic. It can be verified that if κ is even then there is a strategy z that dominates type 0 if and only if $\kappa < 4$. We conclude that if $\kappa \geq 3$, then type 0 is strictly dominated in the type game.

(iii) *Convergence and stability*: After deletion of type 0 we have the payoff matrix $\mathbf{A} = \mathbf{Ic}$ where $\mathbf{b}' = (-1, -b, -1, -b, ...) \in \mathbb{R}^{\kappa}$, and $\mathbf{I}$ is the identity matrix. It is clear that, regardless of whether κ is odd or even, we have $\Pi_1(x) = ... = \Pi_\kappa(x)$ if and only if $x_i = bx_j$, for any odd number $i \leq \kappa$ and any even number $j \leq \kappa$. This is the unique interior Nash equilibrium. In order to show that evolution from any interior initial state converges to a unique interior state it is sufficient to show that the game is stable. Recall that a game with payoff matrix $\mathbf{A}$ is stable if an only if $\mathbf{A}$ is negative definite with respect to the tangent space, i.e. $v \cdot \mathbf{A}v < 0$, for all $v \in \mathbb{R}_0^{\kappa}$, $v \neq \mathbf{0}$. One can transform the problem to one of checking negative definiteness with respect to $\mathbb{R}^{\kappa-1}$ rather than the tangent space $\mathbb{R}_0^{\kappa}$. This is done with the $\kappa \times (\kappa - 1)$ matrix $\mathbf{P}$ (see Weissing (1991)) defined by

$$p_{ij} = \begin{cases} 1 & \text{if } i = j \text{ and } i, j < n \\ 0 & \text{if } i \neq j \text{ and } i, j < n \\ -1 & \text{if } i = n \end{cases} .$$

We need to check whether $(\mathbf{P} \cdot \mathbf{AP})$ is negative definite with respect to $\mathbb{R}^{\kappa-1}$. Let $\mathbf{c}' = (-1, -b, -1, -b, ...) \in \mathbb{R}^{\kappa-1}$. We have $\mathbf{P} \cdot \mathbf{AP} = -(\mathbf{1} \cdot \mathbf{1}') + \mathbf{Ic}$. Note that $v \cdot (-(\mathbf{1} \cdot \mathbf{1}'))\, v < 0$, for all $v \in \mathbb{R}^{\kappa-1}$, $v \neq \mathbf{0}$, so that $v \cdot (\mathbf{P} \cdot \mathbf{AP})\, v < 0$ for all $v \in \mathbb{R}^{\kappa-1}$, $v \neq \mathbf{0}$, if and only if $\mathbf{Ic}$ is negative definite. Since $\mathbf{Ic}$ has two negative eigenvalues $-b$ and -1 it is indeed negative definite. This implies that $\mathbf{A}$ is negative definite with respect to the tangent space. ■

References

Banerjee, A. and Weibull, J. W. (1995), Evolutionary selection and rational behavior, *in* A. Kirman and M. Salmon, eds, 'Learning and Rationality in Economics', Blackwell, Oxford, UK, chapter 12, pp. 343–363.

Basu, K. (1994), 'The travellers' dilemma: Paradoxes of rationality in game theory', *American Economic Review (Papers and Proceedings)* **84**(2), 391–395.

Brown, A., Camerer, C. and Lovallo, D. (2008), To review or not to review? limited strategic thinking at the movie box office. Working paper, Texas AM University.

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

Camerer, C. F., Ho, T.-H. and Chong, J.-K. (2004), 'A cognitive hierarchy model of games', *Quarterly Journal of Economics* **199**, pp. 861–898.

Coricelli, G. and Nagel, R. (2009), 'Neural correlates of depth of strategic reasoning in medial prefrontal cortex', *Proceedings of the National Academy of Sciences USA* **106**(23), 9163–9168.

Costa-Gomes, M. A. and Crawford, V. P. (2006), 'Cognition and behavior in two-person guessing games: an experimental study', *American Economic Review* **96**, 1737–1768.

Crawford, V. (2003), 'Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions', *American Economic Review* **93**, 133–149.

Crawford, V. and Iriberri, N. (2007), 'Level-k auctions: Can boundedly rational strategic thinking explain the winner's curse and overbidding in private-value auctions?', *Econometrica* **75**, 1721–1770.

Dekel, E., Ely, J. C. and Yilankaya, O. (2007), 'Evolution of preferences', *Review of Economic Studies* **74**, 685–704.

Dunbar, R. I. M. (1998), 'The social brain hypothesis', *Evolutionary Anthropology* **6**, 178–190.

Ellingsen, T. and Östling, R. (2010), 'When does communication improve coordination?', *American Economic Review* **100**(4), 1695 – 1724.

Filippov, A. F. (1960), 'Differential equations with discontinuous right-hand side', *Matematicheskii Sbornik* **51**, 199–231. English translation, 1964, American Mathematical Society Translations Series 2, 42, pp. 199-231.

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

Fudenberg, D. and Takahashi, S. (2011), 'Heterogeneous beliefs and local information in stochastic fictitious play', *Games and Economic Behavior* **71**(1), 100–120.

Goeree, J. K. and Holt, C. A. (2004), 'A model of noisy introspection', *Games and Economic Behavior* **46**(2), 365–382.

Haruvy, E. and Stahl, D. O. (2009), Learning transference between dissimilar symmetric normal-form games. Working paper, University of Texas at Austin.

Heller, D. (2004), 'An evolutionary approach to learning in a changing environment', *Journal of Economic Theory* **114**(1), 31–55.

Hofbauer, J. (1995), Stability for the best response dynamics. Working paper, University of Vienna.

Hofbauer, J. and Sigmund, K. (1988), *The Theory of Evolution and Dynamical Systems*, Cambridge University Press, Cambridge.

Holloway, R. (1996), Evolution of the human brain, *in* A. Lock and C. R. Peters, eds, 'Handbook of Human Symbolic Evolution', Clarendon, Oxford, pp. 74–125.

Humphrey, N. K. (1976), The social function of intellect, *in* P. P. G. Bateson and R. A. Hinde, eds, 'Growing Points in Ethology', Cambridge University Press, Cambridge, pp. 303–317.

Ito, T. (1979), 'A filippov solution of a system of differential equations with discontinuous right-hand sides', *Economics Letters* **4**(4), 349–354.

Josephson, J. (2008), 'A numerical analysis of the evolutionary stability of learning rules', *Journal of Economic Dynamics and Control* **32**(5), 1569–1599.

Kinderman, P., Dunbar, R. I. M. and Bentall, R. P. (1998), 'Theory-of-mind deficits and causal attributions', *British Journal of Psychology* **89**, 191–204.

Lahkar, R. and Sandholm, W. H. (2008), 'The projection dynamic and the geometry of population games', *Games and Economic Behavior* **64**(2), 565–590.

Mengel, F. (2009), Learning across games. Working paper, Instituto Valenciano de Investigaciones Económicas.

Mohlin, E. (2011), A note on evolutionary selection against strategies that do not survive iterated elimination of weakly dominated strategies. Manuscript, University College London.

Monderer, D. and Shapley, L. S. (1996), 'Potential games', *Games and Economic Behavior* **14**(1), 124–143.

Nagel, R. (1995), 'Unraveling in guessing games: An experimental study', *American Economic Review* **85**, 1313–1326.

Premack, D. and Wodruff, G. (1979), 'Does the chimpanzee have a theory of mind', *Behavioral and Brain Sciences* **1**, 515–526.

Robson, A. J. (2003), 'The evolution of rationality and the red queen', *Journal of Economic Theory* **111**, 1–22.

Samuelson, L. (2001*a*), 'Analogies, adaptation, and anomalies', *Journal of Economic Theory* **97**(2), 320–366.

Samuelson, L. (2001*b*), 'Introduction to the evolution of preferences', *Journal of Economic Theory* **97**(2), 225–230.

Sandholm, W. H. (2001), 'Potential games with continuous player sets', *Journal of Economic Theory* **97**(1), 81–108.

Sandholm, W. H. (2011), *Population Games and Evolutionary Dynamics*, MIT Press, Cambridge.

Sandholm, W. H. and Dokumaci, E. (2007), 'Dynamo: Phase diagrams for evolutionary dynamics (software suite)', http://www.ssc.wisc.edu/ whs/dynamo.

Sgroi, D. and Zizzo, D. J. (2009), 'Learning to play 3x3 games: Neural networks as bounded-rational players', *Journal of Economic Behavior and Organization* **69**(1), 27–38.

Shapley, L. S. (1964), Some topics in two-person games, *in* M. Dresher, L. S. Shapley and T. A. W., eds, 'Advances in Game Theory', Princeton University Press, Princeton, New Jersey, pp. 1–28.

Stahl, D. O. (1993), 'Evolution of smart$_n$ players', *Games and Economic Behavior* **5**(4), 604–617.

Stahl, D. O. (1999), 'Evidence based rules and learning in symmetric normal-form games', *International Journal of Game Theory* **28**(1), 111–130.

Stahl, D. O. (2000), 'Rule learning in symmetric normal-form games: Theory and evidence', *Games and Economic Behavior* **32**, 105–138.

Stahl, D. O. and Wilson, P. W. (1995), 'On players' models of other players: Theory and experimental evidence', *Games and Economic Behavior* **10**, 218–254.

Steiner, J. and Stewart, C. (2008), 'Contagion through learning', *Theoretical Economics* **3**, 431–458.

Stennek, J. (2000), 'The survival value of assuming others to be rational', *International Journal of Game Theory* **29**, 147–163.

Weibull, J. W. (1995), *Evolutionary Game Theory*, MIT Press, Cambridge Massachusetts.

Weissing, Franz, J. (1991), Evolutionary stability and dynamic stability in a class of evolutionary normal form games, *in* R. Selten, ed., 'Game Equilibrium Models I. Evolution and Game Dynamics', Springer-Verlag, pp. 29–97.

Young, H. P. (1993), 'The evolution of conventions', *Econometrica* **61**(1), 57–84.

Supplement to 'Evolution of Theories of Mind' (*Intended for the web*)

Erik Mohlin*
University College London

May 22, 2011

S1 Stability Concepts

Definition 1 *A closed set $A \subset \Delta(K)$ is* ***Lyapunov stable*** *if every neighborhood (nbd) B of A contains a neighborhood B^0 of A, such that if the system starts in $B^0 \cap \Delta(K)$ at time t_0, then the system remains in B at all times $t \geq t_0$.*

A closed set $A \subset \Delta(K)$ is ***asymptotically stable*** *if it is Lyapunov stable and if there exists a nbd B^* of A such that if the system starts in B^* at t_0 then as $t \to +\infty$ the system goes asymptotically to A.*

The ***basin of attraction*** *of a closed set $A \subset \Delta(K)$ is the set of states such that starting from such a state the system goes to A as $t \to +\infty$.*

A set $A \subset \Delta(K)$ is an ***attractor*** *if its basin of attraction is a nbd of A.*

A strategy $\sigma \in \Delta(S)$ is an ***evolutionarily stable strategy (ESS)*** *if (i) $\tilde{\pi}(\sigma', \sigma) \leq \tilde{\pi}(\sigma, \sigma)$ for all $\sigma' \in \Delta(S)$, and (ii) $\tilde{\pi}(\sigma', \sigma) = \tilde{\pi}(\sigma, \sigma)$ implies $\tilde{\pi}(\sigma', \sigma') < \tilde{\pi}(\sigma, \sigma')$ for all $\sigma' \neq \sigma$.*

Stability of a point is defined as the stability of the singleton $\{x\}$. Note that a Lyapunov stable set is asymptotically stable if and only if it is an attractor. A state is *polymorphic* if it contains positive fractions of more than one type. Otherwise the state is *monomorphic.*

*E-mail: e.mohlin@ucl.ac.uk. Mail: Department of Economics, University College London, Gower Street, London WC1E 6BT, United Kingdom. Telephone: +44 (0)20 7679 5485. Fax: +44 (0)207 916 2775.

S2 Type-Acyclic Games

Lemma S1 *If a game is dominance solvable then it is type-acyclic.*

Proof. First note that $\beta(U)$ survives 1 round of iterated elimination of strictly dominated strategies (IESDS). Next prove that if $\beta^k(U)$ is eliminated by r rounds of IESDS, and if $\beta^{k+1}(U) \neq \beta^k(U)$, then $\beta^{k+1}(U)$ survives $r+1$ rounds of IESDS. To prove this suppose that $\beta^k(U)$ is eliminated by r rounds of IESDS. To obtain a contradiction assume that $\beta^{k+1}(U)$, $\beta^k(U) \neq \beta^{k+1}(U)$ is also eliminated by r rounds of IESDS. This means that $\beta^{k+1}(U)$ earns strictly less than some other strategy s' against all strategies that survive $r-1$ rounds of IESDS. Since $\beta^k(U)$ survives $r-1$ rounds of IESDS this implies that s' earns more than $\beta^{k+1}(U)$ against $\beta^k(U)$, contradicting the definition of β. It follows that $\beta^{k+1}(U)$ survives $r+1$ rounds of IESDS. Finally, an inductive argument yields the conclusion that $\beta^k(U)$ survives at least k rounds of IESDS. Since n is finite iteration of β will lead to the unique Nash equilibrium. ■

Example S1: A type game based on an underlying dominance solvable game may have an asymptotically stable state where $x_\kappa = 0$.

Consider the following game,

	s^1	s^2	s^3	s^4	s^5	s^6	s^7	s^8	s^9
s^1	2	2	2	2	2	2	2	2	2
s^2	30	1	1	1	1	1	1	1	1
s^3	3	3	3	3	3	3	3	3	3
s^4	0	5	0	0	0	0	0	0	0
s^5	0	4	4	4	4	4	4	4	4
s^6	0	-25	0	7	0	0	0	0	0
s^7	0	0	5	5	5	5	5	5	5
s^8	0	0	0	0	0	7	0	7	7
s^9	0	0	0	6	6	6	6	6	6

The game is dominance solvable with the unique Nash equilibrium (s^8, s^8). The type game is as follows,

	$\beta(U) = s^2$	$\beta^2(U) = s^4$	$\beta^3(U) = s^6$	$\beta^4(U) = s^8$
$\beta(U) = s^2$	1	1	1	1
$\beta^2(U) = s^4$	5	0	0	0
$\beta^3(U) = s^6$	-25	7	0	0
$\beta^4(U) = s^8$	0	0	7	7

This game is not dominance solvable. The state where $x_1 = 1/5$ and $x_2 = 1 - x_1$ is asymptotically stable since in that state types 3 and 4 earn $\Pi_3(x) = -25 \cdot \frac{1}{5} + 7 \cdot \frac{4}{5} = \frac{3}{5}$ and $\Pi_4(x) = 0$, respectively, whereas types 1 and 2 earn 1 each.[1]

S3 WBRD-Games

Lemma S2 *If a game is strictly supermodular, then it satisfies WBRD.*

Proof of Lemma S2. Suppose that the game is strictly supermodular. For the setting of symmetric two player games the condition for a game to be strictly supermodular requires that the strategy space can be ordered $S = \{1, 2, ..., n\}$ such that the payoff function exhibits strictly increasing differences, in the sense that if $s > \tilde{s}$ and $s' > \tilde{s}'$ then

$$\pi(s, s') - \pi(\tilde{s}, s') > \pi(s, \tilde{s}') - \pi(\tilde{s}, \tilde{s}').$$

The set of equilibria has some smallest element $s_{\min}^{NE}$ and some largest element $s_{\max}^{NE}$. Vives (1990), theorem 5.1, shows that starting from $s < s_{\min}^{NE}$ ($s > s_{\max}^{NE}$) the Cournot best response dynamic converges monotonically upwards (downwards) to some point in S^{NE}. In a finite game this means that if $s < s_{\min}^{NE}$ then $s < \beta(s) \leq s_{\max}^{NE}$ and there exists some finite k such that $\beta^k(s) \in S^{NE}$. (Similarly, if $s > s_{\max}^{NE}$ then $s > \beta(s) \geq s_{\max}^{NE}$ and there exists some k such that $\beta^k(s) \in S^{NE}$.) Suppose $s < s_{\min}^{NE}$ so that $\beta(s) > s$ (the case with $s > s_{\max}^{NE}$ is exactly parallel). We have $\pi(\beta(s), s) - \pi(s, s) > 0$ and by supermodularity it holds that

$$\pi(\beta(s), s') - \pi(s, s') > \pi(\beta(s), s) - \pi(s, s) > 0,$$

for all $s' > s$. In particular, since $\beta^k(s) > s$ for all $k \geq 1$ it holds that

$$\pi\left(\beta(s), \beta^k(s)\right) - \pi\left(s, \beta^k(s)\right) > 0,$$

for all $k \geq 1$. Finally note that the above inequality is also satisfied for $k = 0$. ■

Example S2: A type game based on an underlying strictly supermodular game may have an asymptotically stable state where $x_\kappa = 0$.

[1] This example also illustrates that if $\kappa < \tilde{k}$ then $\tilde{X}$ need not be asymptotically stable. To see this note that if $\kappa = 3$ then type 3 is strictly dominated in the cognitive game.

Consider the following weakly supermodular game,

$$
\begin{array}{ccccccccc}
 & s^1 & s^2 & s^3 & s^4 & s^5 & s^6 & s^7 & s^8 \\
s^1 & 3 & 1 & 1 & -\frac{1}{10} & 1 & 1 & 1 & 1 \\
s^2 & 3 & 1 & 1 & -\frac{1}{10} & 1 & 1 & 1 & 1 \\
s^3 & 3 & 1 & 1 & -\frac{1}{10} & 1 & 1 & 1 & 1 \\
s^4 & \frac{31}{10} & \frac{11}{10} & \frac{11}{10} & 0 & \frac{11}{10} & \frac{11}{10} & \frac{11}{10} & \frac{11}{10} \\
s^5 & -3 & 1 & 1 & 1 & \frac{21}{10} & \frac{21}{10} & \frac{21}{10} & \frac{21}{10} \\
s^6 & -10 & -2 & 0 & \frac{13}{10} & \frac{24}{10} & \frac{24}{10} & \frac{24}{10} & \frac{24}{10} \\
s^7 & -11 & -3 & -\frac{3}{10} & 1 & 3 & 3 & 3 & 3 \\
s^8 & -12 & -4 & -\frac{3}{10} & 1 & 3 & \frac{31}{10} & \frac{31}{10} & \frac{31}{10}
\end{array}.
$$

The type game is as follows,

$$
\begin{array}{ccccc}
 & U & \beta(U) = s^4 & \beta^2(U) = s^6 & \beta^3(U) = s^8 \\
U & \frac{197}{245} & \frac{4}{7} & \frac{21}{10} & \frac{21}{10} \\
\beta(U) = s^4 & \frac{97}{70} & 0 & \frac{11}{10} & \frac{11}{10} \\
(\beta(U))^2 = s^6 & -\frac{11}{70} & \frac{13}{10} & \frac{24}{10} & \frac{24}{10} \\
(\beta(U))^3 = s^8 & -\frac{3}{7} & 1 & \frac{31}{10} & \frac{31}{10}
\end{array}.
$$

Consider a state where $x_1 = 1 - x_0$ and type 0 and 1 earn the same. For this state to be a rest point it must solve $\frac{197}{245}x_0 + \frac{4}{7}(1 - x_0) = \frac{97}{70}x_0$, which is done by the state x^* where $x_0^* = \frac{56}{113}$ and $x_1^* = 1 - x_0^*$. If we restrict attention the type game with only type 0 and type 1 then the state with $x_0 = \frac{56}{113}$ and $x_1 = 1 - x_0$ is asymptotically stable (notice the hawk-dove character of this submatrix). Moreover, if we again consider the full type game, then type 1 and 2 earn $\Pi_0(x) = \Pi_1(x) = \frac{97}{70} \cdot \frac{56}{113} = \frac{388}{565} \approx 0.686\,73$, in state x^*. The payoffs to type 2 and 3 are $\Pi_2(x) = -\frac{11}{70}\left(\frac{56}{113}\right) + \frac{13}{10}\left(1 - \frac{56}{113}\right) = \frac{653}{1130} \approx 0.577\,88$, and $\Pi_3(x) = -\frac{3}{7}\left(\frac{56}{113}\right) + \left(1 - \frac{56}{113}\right) = \frac{33}{113} \approx 0.292\,04$, respectively in state x^*. Thus the state x^* is asymptotically stable. The game defined by the above matrix is only weakly supermodular. To obtain a strictly supermodular game add a factor $\varepsilon \cdot i \cdot j$ to each entry ij in the matrix. For sufficiently small $\varepsilon \in \mathbb{R}_+$ the type game based on such a strictly supermodular game will have the same stability properties as the game with $\varepsilon = 0$.

The proof of proposition 2 refers to a result in Mohlin (2011), a specialized version of which is provided here.

Lemma S3 *In the type game based on an underlying WBRD-game played by level-k types $K \backslash \{0\}$, with $\kappa = k^{NE}$, evolution from any interior initial condition converges to the state where $x_\kappa = 1$.*

Proof. The payoffs of the type game are described in the proof of proposition 2. We proceed with a proof by induction. Fix k and let $\beta = w(k+1,k) - w(k,k) > 0$, and $\gamma = \max_{i\leq k-1}[|w(k,i) - w(k+1,i)|] \geq 0$. It follows that

$$\Pi_{k+1}(x) - \Pi_k(x) \geq \beta x_k - \gamma \sum_{i=1}^{k-1} \delta x_i. \tag{1}$$

Inductive hypothesis: Consider $k \leq \kappa - 1$. As inductive hypothesis suppose that (a) for all $i \in \{1, ..., k-1\}$ it holds that $\xi_i(t, x^0) \to 0$, and (b) for some $I \subseteq \{1, ..., k-1\}$ it holds that

$$\lim_{t\to\infty} \frac{\sum_{i\in I} \xi_i(t, x^0)}{\xi_k(t, x^0)} = 0, \tag{2}$$

and (c) for all $i \in L = \{1, ..., k-1\} \backslash I$ it holds that $\int_0^t \xi_i(\tau, x^0)\, d\tau \to \alpha_i$ for some finite $\alpha_i \in \mathbb{R}_+$.

Inductive step: Define $v_k(x) = \log x_k - \log x_{k+1}$. The derivative with respect to time is, using (1),

$$\dot{v}_k(x) = \frac{\dot{x}_k}{x_k} - \frac{\dot{x}_{k+1}}{x_{k+1}} = \Pi_k(x) - \Pi_{k+1}(x) \leq \sum_{i=1}^{k-1} \gamma x_i - \beta x_k.$$

Integration w.r.t. t yields

$$\begin{aligned} v_k\left(\xi\left(t, x^0\right)\right) &\leq v_k\left(x^0\right) + \sum_{i\in L} \gamma\alpha_i + \sum_{i\in I} \gamma \int_0^t \xi_i\left(\tau, x^0\right) d\tau - \beta \int_0^t \xi_k\left(\tau, x^0\right) d\tau \\ &= v_k\left(x^0\right) + \sum_{i\in L} \gamma\alpha_i + \int_0^t \sum_{i\in I} \gamma\xi_i\left(\tau, x^0\right) d\tau - \int_0^t \beta\xi_k\left(\tau, x^0\right) d\tau \end{aligned} \tag{3}$$

Note that the integral $\int_0^t \xi_k(\tau, x^0)\, d\tau$ is increasing in t so either (case 1) it goes to $+\infty$ or (case 2) it goes to some finite $\alpha_k \in \mathbb{R}_+$. It follows from (2) that there is some t' such that if $t > t'$ then

$$\sum_{i\in I} \gamma\xi_i\left(t, x^0\right) \leq \beta\xi_k\left(t, x^0\right). \tag{4}$$

Case 1: If $\int_0^t \xi_k(\tau, x^0)\, d\tau \to \infty$ then (3) and (4) together imply $v_k(\xi(t, x^0)) \to -\infty$ so that the definition of v_k implies $\xi_k(t, x^0) \to 0$ and $\xi_k(t, x^0) / \xi_{k+1}(t, x^0) \to 0$.

Together with (2) this further implies

$$\lim_{t\to\infty} \frac{\sum_{i\in I} \xi_i(t, x^0) + \xi_k(t, x^0)}{\xi_{k+1}(t, x^0)} = 0.$$

Case 2: If instead $\int_0^t \xi_k(\tau, x^0)\, d\tau \to \alpha_k < \infty$ then $\xi_k(t, x^0) \to 0$, by the fact that $\xi_k(\tau, x^0)$ is uniformly continuous in t. To see that $\xi_k(\tau, x^0)$ is uniformly continuous in t note that $\left|\frac{\partial}{\partial t}(\xi_k(t, x^0))\right| \le \max_{x\in\Delta(K)} \left|\Pi_k(x) - \bar{\Pi}(x)\right| \in \mathbb{R}_+$. (For an argument why uniform continuity implies $\xi_k(t, x^0) \to 0$, see Weibull (1995), proposition 3.2.) Moreover, if $\int_0^t \xi_k(\tau, x^0)\, d\tau \to \alpha_k < \infty$ then it follows from (2) that $\lim_{t\to\infty} \sum_{i\in I} \xi_i(t, x^0) < \infty$, and hence for all $i \in I$, we have $\xi_i(t, x^0) \to \alpha_i < \infty$, for some α_i.

Inductive base case: Define $v_1(x) = \log x_1 - \log x_2$. The derivative with respect to time is

$$\dot{v}_1(x) = \frac{\dot{x}_1}{x_1} - \frac{\dot{x}_2}{x_2} \le -\beta_1 x_1,$$

where $\beta_1 = w(2,1) - w(1,1) > 0$. Integrating both sides w.r.t. t yields $v_1(\xi_1(t, x^0)) \le v_1(x^0) - \beta_1 \int_0^t \xi_1(\tau, x^0)\, d\tau$. If $\int_0^t \xi_1(\tau, x^0) \to +\infty$ then by the definition of $v_1(x)$ we have $\xi_1(t, x^0) \to 0$ and $\xi_1(t, x^0)/\xi_2(t, x^0) \to 0$. If $\int_0^t \xi_1(\tau, x^0)\, d\tau \to \alpha_1$ then $\xi_1(\tau, x^0) \to 0$ by the fact that $\xi_1(\tau, x^0)$ is uniformly continuous in t. ■

References

Mohlin, E. (2011), A note on evolutionary selection against strategies that do not survive iterated elimination of weakly dominated strategies. Manuscript, University College London.

Vives, X. (1990), 'Nash equilibrium with strategic complementarities', *Journal of Mathematical Economics* **19**, 305–321.

Weibull, J. W. (1995), *Evolutionary Game Theory*, MIT Press, Cambridge Massachusetts.