Maastricht University

Abhimanyu Khan, Ronald Peeters

Cognitive hierarchies in adaptive play

RM/12/007

METEOR

Cognitive hierarchies in adaptive play[*]

Abhimanyu Khan[†]   Ronald Peeters[‡]

January 2012

Abstract

Inspired by the behavior in repeated guessing game experiments, we study adaptive play by populations containing individuals that reason with different levels of cognition. Individuals play a higher order best response to samples from the empirical data on the history of play, where the order of best response is determined by their exogenously given level of cognition. As in Young's model of adaptive play, (unperturbed) play still converges to a minimal curb set. However, with the random perturbations of this (higher order) best response dynamic, the stochastically stable states obtained may now be different, but in a deterministic manner. Perhaps counter-intuitively, higher cognition may actually be bad for both the individual with higher cognition and his parent population.

Keywords: evolution of behavior, adaptive play, cognitive hierarchies, level-k reasoning.

JEL classification codes: C73, D03.

[†]Department of Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Email: `a.khan@maastrichtuniversity.nl`.

[‡]Corresponding author. Department of Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Email: `r.peeters@maastrichtuniversity.nl`.

1 Introduction

One of the primary focuses of game theory is to model individual decision making in settings of strategic conflict. The more traditional workhorse employed to this end has been the assumption of rationality and common belief of it, which at times might be a very demanding requirement. Binmore (1987) discusses fundamental issues with this assumption, including the inconsistency that it cannot be replicated by a Turing machine. Furthermore, experiments on human subjects reveal that for certain applications, the assumptions of common belief of rationality and its associated implication of an infinite depth of reasoning appear too strong, resulting in a disparity between theoretical predictions and empirical observations in the laboratory. This has generated considerable interest in a paradigm where individuals are either capable of, or make use of finite depth of reasoning. This could be either because of cognitive limitations on individuals or because of a belief that the co-player in the strategic situation will employ finite depth of reasoning, in which case it is also optimal to do so.

To illustrate this, consider the guessing game experiment of Nagel (1995) wherein individuals had to guess the number in the interval $[0, 100]$ which would be the closest to $p < 1$ times the mean of all guesses. Under common belief of rationality, the only rationalizable choice is to guess 0. The empirical data of the experiment, though, suggests that only a handful of the subjects guess 0. A considerable amount of the chosen strategies can be described by what is called "level-k" behavior. Here, level-1 behavior refers to best responding to the belief that mean of the guesses is going to be clustered around 50. This could be either due to the salience of the number 50 (Schelling salience) or because of the belief that the guesses of the others can be approximated by a uniform distribution over the interval $[0, 100]$. In either case, the optimal guess arising out of level-1 behavior is $50\,p$. Level-2 behavior supposes that the others will employ level-1 reasoning and guess $50\,p$. Consequently, it is optimal to guess $50\,p^2$. Higher order level-k behavior is defined iteratively with level-k behavior being identified with guessing $50\,p^k$. Nagel (1995) finds that most of the subjects exhibit level-1 and level-2 behavior. The decisions in subsequent repetitions show a declining trend in the guesses. Two mutually non-exclusive explanations are an increase in the subjects' depth of reasoning and a shift in reference point for the most primary belief. Nagel's (1995) data does not provide any evidence for the former explanation and leans strongly towards the second explanation: subjects seem to use the mean of the previous period as the reference point.

A fair amount of theoretical and experimental studies has been directed towards the development of models that explicitly take into account the beliefs of individuals about other individuals' strategic decisions. For example, Stahl (1993) and Stahl and Wilson (1995) present a theoretical model of such players. The latter use a series of experiments to test for this pattern. Camerer et al. (2004) develop a cognitive hierarchy model in which each player assumes that she is the most sophisticated type and that the other individuals are of lower

cognition. These models seem to explain the pattern of play in initial rounds of experiments – such as the one of Nagel (1995) – fairly well. They also are found to be robust to a wide specification of games: cf. Stahl and Wilson (1994, 1995), Crawford and Iriberri (2007), Wang et al. (2010) and Coricelli and Nagel (2009).

Our interest, on the other hand, is to examine the effect of such a hierarchical model of players on the long-run process. Thus, not only are we interested in a framework where we begin with the hypothesis that individuals come in varying degrees of sophistication, but allow the different levels of cognition to operate on past experience as indicated by the empirical data of play. While the sophistication of a particular individual does not change over time, individuals pay attention to the way in which the game has unfolded in the recent past. The importance of modeling this feature is also emphasized in Binmore (1988). Although his main focus is on, what he calls, the "eductive process" (i.e., the cognitive process of the individual during the decision-making) rather than the "evolutive process", he recognizes the importance of the latter as well. In this paper, we consider a process which, while evolutionary in nature, also has the flavor of the eductive process. We believe that Young's (1993a, 1998) model of adaptive play provides a very convenient platform to incorporate these elements.

In Young's model of adaptive play, the interaction is modeled as a population game, where a population of players is associated with a specific role in a game. Populations have, at their disposal, a record of recent history of play. In each period, an individual is drawn randomly from the parent population to play the game. The individual samples incompletely and without replacement from the past play of the rival population and plays a best response to the empirical distribution of strategies as revealed by the sample. In terms of the level-k model, this means that the individuals, by best responding to the empirical distribution available to them, exhibit level-1 behavior. Using an argument similar to Hurkens (1995), Young (1998) demonstrates that play converges almost surely to a minimal curb set.[1] The minimal curb sets form the recurrent classes (absorbing sets) of the dynamic process described by simple best response behavior. In order to select amongst the minimal curb sets (which may be numerous), the notion of stochastic stability is adopted.

We introduce the element of sophistication in this model of adaptive play in order to study the effect of higher cognition on long run outcomes. Young's model, with only simple (best responding) players, serves as the point of reference. We postulate that populations are comprised of individuals of varying levels of sophistication.[2] The most primitive behavioral trait is described by the level-1 individuals referenced to above.[3] A level-2 individual holds the

---

[1]The minimal curb set is, loosely speaking, a set which contains all its best responses and there is no proper subset contained in it with the same property. This concept was first introduced by Basu and Weibull (1991).

[2]Mohlin (forthcoming) shows that it is possible for evolutionary learning processes to converge to a state where different cognitive types co-exist.

[3]For the purpose of nomenclature, we retain the associated terminology of the level-k model, even though we step outside the boundaries of it. Our focus is the long-run behavior, whereas the level-k model is meant for the purpose of explaining initial play.

belief that the rival is a level-1 type. Consequently, she samples from the actions of the own population to estimate the best response of a level-1 individual in the rival population and best responds to the estimate just formed. A level-3 individual plays a best response to the belief that the rival is of level-2. Hence, the level-3 individual samples from the actions of the rival population to estimate the best response of a level-1 individual in the own population. This is then used to estimate the best response of a level-2 individual in the rival population. The level-3 individual plays a best response to this estimate of the strategy that might be used by the level-2 individual of the rival population. This process is iterated one step more for each higher level of sophistication. So, as in Nagel (1995), each player assumes that the other players are one step down on the cognitive scale.

In Proposition 1, we show that the convergence to a minimal curb set, as shown by Young (1998) for the setting with only simple players, is invariant to the introduction of any higher level of sophistication in any population. A refinement of the minimal curb sets of the game may be obtained by considering an infinitesimally small probability of mistakes or experimentation, whereby individuals choose any strategy (and not necessarily a best response of some order). This selects the set of stochastically stable states. The possibility of mistakes or experimentation now makes transition across minimal curb sets possible and the minimal curb sets that are most "stable" to such mistakes comprise the set of stochastically stable states. In Propositions 2–4, we find that the stochastically stable set may be sensitive to the composition of cognitive types in the populations. Proposition 2 shows this for the specific instance of the evolutionary Nash bargaining game while Propositions 3 and 4 do so for generic games.

In Young (1993b) it is shown that in the evolutionary Nash bargaining game with simple players only, the long-run outcome is the generalized Nash bargaining solution with the bargaining power being proportional to the populations' own sample size. The intuition is simple: the smaller the sample size, the higher is the responsiveness to mistakes in the other population, leading to a smaller bargaining power.[4] In Proposition 2 we show that the introduction of more sophisticated individuals in a single population has no effect on the long-run outcome in case this population has the larger sample size, but that the standard Nash bargaining solution (that is, equal bargaining power) is obtained as long-run outcome in case this population has a smaller sample size. This effect has already been shown by Sáez-Martí and Weibull (1999) when allowing for "clever" individuals (that is, of level-2) on one side. They conclude that a grain of cleverness compensates for the lower bargaining power from smaller sample size but cleverness confers no additional benefit apart from that. This finding is driven by the fact that level-2 individuals best respond to a best response to the sample of their own population. Via their believes, the level-2 individuals, when endowed with a smaller

---

[4]Although in the Nash bargaining game a higher sample size confers a benefit to the population, depending on the payoff structure it can be a bane as well – see Section 6 for more details.

sample size, make the rival population appear more responsive to mistakes made by the own population. We show that there are no benefits of adding individuals of even higher level of cognition. Moreover, we show that if both populations host sophisticated individuals, then we obtain the standard Nash bargaining solution as the stochastically stable state.

In Proposition 3 we show that for generic games, the introduction of a fraction of sophisticated individuals in one population has an effect on the stochastically stable set only if the population with these individuals has a smaller sample size. Proposition 4 claims that if both populations host individuals with at least level-2 cognition, then irrespective of the relative sample size, the stochastically stable set is the same as that in Young's adaptive play model with both populations having equal sample size. Thus, the effect of higher cognition is to "decrease" the sample size of the rival population to that of the own population. Matros (2003) already showed this for the situation where one population has level-2 individuals and both populations have the same sample size. Our result extends Matros (2003), as we allow for even higher levels of cognition, populations to have unequal sample sizes, and both populations to host individuals of level-2 or higher.

In the following section, we specify the model of adaptive play with higher cognitive individuals in more detail and state that this process of adaptive play converges to a minimal curb set. In Section 3 and Section 4, we study the impact of higher cognitive types on the long-run outcome for the evolutionary Nash bargaining game and general games respectively. In Section 5, we provide a discussion on two possible alternative specifications of our model. In our model we assume that an individual of a particular level believes she is playing someone of one level lower. In the first alternative specification, we show that our results are robust to the alternative specification where she believes her rival to be of any lower level – as is assumed for instance in Stahl (1993) and Camerer et al. (2004). In the second alternative specification, for a specific class of games, we relax the knowledge (individuals within) populations may have about cardinal preferences of (individuals within) the rival population. As a result, higher cognitive types use their own utility function while forming a belief about the rival's play. Binmore (1988) mentions that such form of introspection might be a plausible way in which higher cognition operates. We find that play still converges to a minimal curb set (as ordinal preferences do not change), but the stochastically stable set may change (as the sensitivity to mistakes may change): in the evolutionary Nash bargaining game, the generalized Nash bargaining solution with bargaining power proportional to the rivals' sample size and the 50–50 split may become feasible long-run outcomes. Finally, we close with a discussion of the implication of our results in Section 6. Amongst other things, we discuss the qualitative impact of higher cognition. The distribution of strategies in the long-run outcome might be affected, possibly resulting in payoff differences across individuals within and across populations. Further it is, perhaps counter-intuitively, shown that an entire population can be worse-off by the presence of more cognitive individuals.

2   Preliminaries

There are two finite populations A and B and each population is assigned a specific role in a specified two-player game that is played recurrently between randomly selected individuals from the two populations. Each individual has a positive probability of representing the population that she belongs to.[5] The selected individual from population A (B) has to choose a strategy x_A (x_B) out of a finite set of strategies X_A (X_B). The pair of chosen strategies $(x_A, x_B) \in X_A \times X_B$ yields a payoff of $\pi_A(x_A, x_B)$ and $\pi_B(x_A, x_B)$ to the individual from population A and B respectively. We assume each population to be homogenous, i.e. all individuals in a population have the same utility function. The only heterogeneity we assume within a population is in the level of sophistication.

We assume individuals play adaptively. Both populations have access to the strategies chosen in the last m periods of play and individuals in population A (B) can draw a sample of proportion a (b) – i.e., of length $a\,m$ ($b\,m$) – without replacement. We assume each possible sample to have a positive probability of being drawn, but do not require all samples in the history to be equally likely to be drawn.

We identify the empirical distribution over chosen strategies with level-0 ($L0$) behavior. So, a level-1 ($L1$) individual plays a best response to her sample from the rival population's strategies.[6] Next, a level-2 ($L2$) individual holds the belief that her rival is $L1$ and attempts to best respond to the strategy that the $L1$ individual of the rival population might choose. As a result, the $L2$ individual samples from her own population's past play, formulates the best response to the drawn sample[7], adopts this as estimate of the strategy that the rival $L1$ individual might use, and plays a best response to it. The behavior of any individual of higher level of sophistication is described analogously.

It is evident from the above that an Lk individual, with k odd, draws a sample from the strategies of the rival population. In contrast, an Lk individual, with k even, draws a sample from the strategies of the own population. Following that, the sample is processed in accordance to the cognition of the individual.

At this point, we make the assumption that if a population contains an Lk individual with $k \geq 2$, then it also contains an $L(k-1)$ individual.[8] So, we do not allow the entire population to be be composed of, say, $L4$ individuals only. And, in case a population contains an $L4$ individual, then it also contains at least one $L3$, $L2$ and $L1$ individual. This permits us to describe the cognitive types present in a population by the most sophisticated individual. We use $\hat{k}_A$ and $\hat{k}_B$ to refer to the highest cognitive level in population A and population B

---

[5]We do not require the probability to be selected to be equal for each individual in a population.

[6]In cases of multiple best responses, we always assume each best response to be chosen with positive probability, not necessarily with equal chance.

[7]In order to do so, it is necessary that the $L2$ individual possess knowledge of the utility function of the rival population.

[8]In Section 5 we replace this assumption with an alternative one.

respectively.

Given any arbitrary history on m periods of play, the adaptive process described above yields a Markov process on the state space $\Omega = (X_A \times X_B)^m$, where the states are the possible histories of length m. Let this (unperturbed) process, for sample proportions a and b and highest levels of sophistication $\hat{k}_A$ and $\hat{k}_B$, be denoted by $P^{m,a,\hat{k}_A,b,\hat{k}_B}(0)$.

Let C_i be a nonempty subset of X_i ($i = A, B$). We denote the set of probability distributions over C_i by $\Delta(C_i)$. Moreover, by $BR_i(C_j)$, we denote the set of strategies in X_i that consists of the best replies to any mixture in $\Delta(C_j)$ for the individuals in population i ($j \neq i$). Now, we can define the notion of (minimal) curb sets, which is due to Basu and Weibull (1991). The product set $C = C_A \times C_B$ is *closed under best replies* (or C is a *curb set*) if $BR_A(C_B) \times BR_B(C_A) \subseteq C$. Such a curb set is *minimal* if it does not properly contain a curb set.

Proposition 1. *If the history m is sufficiently large and the sampling is sufficiently incomplete (i.e., a and b are sufficiently small), then the minimal curb sets are the recurrent classes of the process of adaptive play $P^{m,a,\hat{k}_A,b,\hat{k}_B}(0)$.*

Proof. The proof for this is essentially the same as in Young (1998, Thm. 7) and Matros (2003, Thm. 1). Young (1998) shows that best response adaptive play converges to a minimal curb set in a finite number of steps when there are only $L1$ individuals in either population. Since in our setting there is for any finite number of periods, a positive probability that only $L1$ individuals are chosen during these periods, we know that our process converges to a minimal curb set. Matros (2003) shows that a fraction of $L2$ individuals in one population does not have any disruptive effect on the convergence to a minimal curb set. The idea is that, by definition of the minimal curb set, higher order best responses are contained in it. Using the same arguments, we can show that the same holds for Lk individuals in either population. After play converges to a minimal curb set (by successively drawing $L1$ individuals), responses of all higher cognitive types are contained in it such that play never leaves it. Furthermore, it is possible to transit from one state in the recurrent class to another in finite time by considering the event that only $L1$ individuals are chosen. The result then follows from Young (1998). ∎

The model of adaptive play so far has been built on the assumption that individuals are simple or sophisticated best responders. This results in the convergence to a minimal curb set. So, minimal curb sets represent the recurrent classes of the (unperturbed) Markov process. In order to allow for transits from one minimal curb set to another, we need to perturb the process. For the perturbed process, we assume that each individual has a probability ε to experiment or to commit a mistake; that is, they may choose any strategy – even those that are not a best response or a best response of a best response and so on, to any conceivable sample drawn from the history of past play. These experimentations induce the

resulting perturbed process of adaptive play $P^{m,a,\hat{k}_A,b,\hat{k}_B}(\varepsilon)$ to be ergodic, such that for each $\varepsilon > 0$ a unique stationary distribution exists; that is, a unique solution $\mu(\varepsilon)$ to the equation $\mu \cdot P^{m,a,\hat{k}_A,b,\hat{k}_B}(\varepsilon) = \mu$. The *stochastically stable set* is defined to consist of precisely those states that receive a positive weight in the limiting stationary distribution $\mu^* = \lim_{\varepsilon \downarrow 0} \mu(\varepsilon)$.

The stochastically stable set is a subset of the set of recurrent classes (or, in the present case, the minimal curb sets) of the unperturbed adaptive process $P^{m,a,\hat{k}_A,b,\hat{k}_B}(0)$. Intuitively, this set consists of the minimal curb sets that are easier to reach and more difficult to transit from with (a series of) experimentations. We now make an important observation. We are interested in the minimum number of experimentations that thereafter make a transition from one minimal curb set to another possible via the best-response dynamic. The best responses of Lk individuals, with k odd (even), are only affected by a sufficient number of experimentations in the rival (own) population. For higher cognitive types, an Lk individual will have a best response outside the prevailing minimal curb set only if she believes that the $L(k-1)$ individual in the rival population will do so. But this happens if she believes that the $L(k-1)$ individual believes that an $L(k-2)$ individual will do so. Ultimately, this can be traced to the belief that an $L1$ individual in some population chooses a strategy outside the recurrent class. If this does not happen, then none of the higher cognitive individuals in any population believe that the actions of the rival is going to change. The belief that an $L1$ individual in the rival (own) population does so, induces a change in behavior by Lk individuals, with k odd (even).

3   The evolutionary Nash bargaining game

The Nash bargaining game is a two-player game where each of the two players demands a portion of some good (or, amount of money). If the total amount requested by the players is less than that available, both players get their request. If their total request is greater than that available, neither player gets their request. In this section we consider the Nash bargaining game in an evolutionary framework, played recurrently by randomly selected players from two finite populations.

Let A and B be two populations consisting of a finite number of individuals. In each period t, one individual is selected at random from each population to play the Nash bargaining game. Simultaneously and independently, each individual announces a demand from the feasible set of demands $D(\delta) = \{\delta, 2\delta, ..., 1 - \delta\}$, say x_A^t and x_B^t. Each of them receives their respective demand if the two demands sum up to not more than the whole (that we normalize to one); i.e, if $x_A^t + x_B^t \leq 1$. Otherwise, both receive nothing. We assume that all individuals in population A have the same concave, strictly increasing, differentiable von Neumann-Morgenstern utility function $u : [0, 1] \rightarrow \mathbb{R}$ with $u(0) = 0$. Similarly, we assume all individuals in population B to have a utility function $v : [0, 1] \rightarrow \mathbb{R}$ with the same properties.

At period t, individuals α and β are chosen from populations A and B. Both individuals have access to a record of play of the last m periods: $\omega^t = ((x_A^{t-m}, x_B^{t-m}), \ldots, (x_A^{t-1}, x_B^{t-1})) \in \Omega = (D(\delta) \times D(\delta))^m$. Individual α can draw a sample of size $a\,m$ of demand pairs (x_A^τ, x_B^τ) from the last m periods of play without replacement. Similarly, individual β draws a sample of size $b\,m$ of demand pairs (x_A^τ, x_B^τ) from the last m periods of play without replacement. We do not require each pair to be drawn with equal probability, but it is essential that any sample of the appropriate size can be drawn. Next, individual α (β) makes a demand x_A^t (x_B^t) that maximizes her expected payoff against the empirical distribution of demands as given by the sample drawn (given her depth of reasoning).

A state $\omega \in \Omega$ is a *convention* if it consists of some fixed division $(\bar{x}_A, \bar{x}_B) \in D(\delta) \times D(\delta)$ repeated m times in succession and we call it an *efficient convention* if, in addition, $\bar{x}_A + \bar{x}_B = 1$. In Theorem 1 of Young (1993b) it is shown – for populations solely consisting of $L1$ individuals – that if at least one individual in each population samples at most half of the record of play, then (from any initial state) the process converges almost surely to an efficient convention. This is actually a corollary of Proposition 1, as the set of efficient conventions coincides with the set of all minimal curb sets. In the remainder of this section, we will focus on the (set of) stochastically stable state(s) as the precision of demand δ goes to zero.

In Young (1993b) it is shown that when both populations consist solely of $L1$ individuals ($\hat{k}_A = \hat{k}_B = 1$), the stochastically stable state is the generalized Nash bargaining solution with bargaining power proportional to the sample sizes, which is the efficient convention given by the division $(\bar{x}_A, \bar{x}_B)$ that maximizes the Nash product $(u(x_A))^a (v(x_B))^b$ subject to $x_A, x_B \in [0, 1]$ and $x_A + x_B = 1$.

Sáez-Martí and Weibull (1999) introduce, in one of the populations, clever agents ($L2$ individuals) who possess knowledge of the utility function of the members of the rival population. These clever players estimate the best reply of the opponent by sampling from the record of their own population and then best reply to this estimate. Let, without loss of generality, these clever individuals be members of population A, such that $\hat{k}_A = 2$ and $\hat{k}_B = 1$. It is demonstrated that if the proportion of these clever agents is less than 1, then the stochastically stable division – as the precision of demand δ tends to zero – is the generalized Nash bargaining solution with the bargaining power proportional to the sample sizes a and b if $a > b$, and to the standard Nash bargaining solution if $a \leq b$. So, interpreting the sample size in Young's model as a measure of the bargaining power of a population, cleverness in a population has the effect of compensating for lower bargaining power arising due to smaller sample size, but it does not have an effect if the clever population has a larger sample size (and hence higher bargaining power).

We relax the restriction on the level of sophistication of the individuals in the populations in two directions. First, we allow the highest level of cognition in population A to be larger than two: $\hat{k}_A > 2$. Second, we allow the possibility of clever (and even more sophisticated)

agents in the B population as well: $\hat{k}_B \geq 2$.

We find that if population B has only $L1$ individuals while population A has Lk individuals with $k > 2$ (i.e., $\hat{k}_A > 2$ and $\hat{k}_B = 1$), then nothing changes with respect to the result of Sáez-Martí and Weibull (1999). However, when population B also holds individuals of level of sophistication larger than one (i.e., $\hat{k}_A, \hat{k}_B \geq 2$), then the standard Nash bargaining solution is the unique stochastically stable state.

We summarize all results in the following proposition, where part (i) is due to Young (1993b) and part (ii) is demonstrated for $L2$ individuals in one population by Sáez-Martí and Weibull (1999).

Proposition 2. *(i) Suppose that both populations comprise of $L1$ individuals. The stochastically stable state is then the generalized Nash bargaining solution, with the bargaining power of each population equal to the sample proportion of each population.*

(ii) Suppose population A has a fraction of sophisticated best responders. Then, the stochastically stable state is the generalized Nash bargaining solution if $a > b$. Otherwise, the stochastically stable state is the standard Nash bargaining solution.

(iii) Suppose both populations have sophisticated best responders. The stochastically stable state is the standard Nash bargaining solution.

Proof. With unperturbed adaptive play, the play converges to an efficient convention. This is because the minimal curb sets of the games coincide with the efficient conventions. To identify the stochastically stable state, we need to look at the minimum number of experimentations needed to transit from one convention to another.

Part (i). When there are only $L1$ individuals in each population, then the minimum number of mistakes needed to move away from the efficient convention $(x, 1-x)$ corresponds to the smallest integer greater than or equal to $m\, r^\delta(x)$, where

$$r_A^\delta(x) = a\, \min\{\tfrac{u(x)-u(x-\delta)}{u(x)}, \tfrac{u(x)}{u(1-\delta)}\},\tag{1}$$

$$r_B^\delta(x) = b\, \min\{\tfrac{v(1-x)-v(1-x-\delta)}{v(1-x)}, \tfrac{v(1-x)}{v(1-\delta)}\},\ \text{and}\tag{2}$$

$$r^\delta(x) = \min\{r_A^\delta(x), r_B^\delta(x)\}.\tag{3}$$

To explain this, note that there are two ways that we can move away from a convention. One is when individuals from population B make mistakes and these mistakes are sampled by individuals in population A. The other is when individuals from population A make mistakes and these mistakes are sampled by individuals in population B. Equation (1) captures the former possibility; Equation (2) the latter.

There are two mistakes an individual of population B can make: demanding an amount more than $1-x$ and demanding an amount less than $1-x$. Amongst the demands greater than $1-x$, the least resistant one is to ask for δ more: i.e., demanding $1-x+\delta$. A best responding individual of population A would shift to asking $x - \delta$, when this mistake has been made

sufficiently frequent. Let ℓ represent the least number of times population B has to make this mistake in order to displace the convention. Given that population A individuals draw a sample of size $a\,m$, they may shift to asking $x - \delta$ when $\frac{a\,m - \ell}{a\,m}\,u(x) \leq u(x - \delta)$. This gives the least proportion of mistakes $\frac{\ell}{m} = a\,\frac{u(x) - u(x - \delta)}{u(x)}$, which is one of the terms in Equation (1). The second term comes from the situation when population B demands less than $1 - x$ by mistake. Among these mistakes, demanding δ is the least resistant one. Equation (2) can be derived in a similar manner.

Consider the fraction

$$\frac{r^\delta(x)}{\delta} = \min\{a\,\frac{u(x) - u(x - \delta)}{\delta\,u(x)}, a\,\frac{u(x)}{\delta\,u(1 - \delta)}, b\,\frac{v(1 - x) - v(1 - x - \delta)}{\delta\,v(1 - x)}, b\,\frac{v(1 - x)}{\delta\,v(1 - \delta)}\}. \tag{4}$$

As $\delta \downarrow 0$, the first term can be written as $a\,\frac{u'(x)}{u(x)}$ and the third as $b\,\frac{v'(1 - x)}{v(1 - x)}$, while the second and fourth terms become unbounded. So, $\lim_{\delta \downarrow 0} \frac{r^\delta(x)}{\delta} = \min\{a\,\frac{u'(x)}{u(x)}, b\,\frac{v'(1 - x)}{v(1 - x)}\}$. The efficient convention with the highest minimum resistance is the efficient convention $(\bar{x}, 1 - \bar{x})$ with $\bar{x}$ maximizing the latter expression. As $\frac{u'(x)}{u(x)}$ is decreasing in x and $\frac{v'(1 - x)}{v(1 - x)}$ is increasing in x, the maximum is at the (unique) solution to $a\frac{u'(x)}{u(x)} = b\frac{v'(1 - x)}{v(1 - x)}$. This is precisely the first-order condition for the maximization of $(u(x))^a(v(1 - x))^b$; i.e., the generalized Nash bargaining solution.

Part (ii). When population A contains Lk individuals with $k \geq 2$, while population B only contains $L1$ individuals, then the minimum number of mistakes needed to move away from the efficient convention $(x, 1 - x)$ corresponds to the smallest integer greater than or equal to $m\,r^\delta(x)$, where

$$r_A^\delta(x) = a\,\min\{\frac{u(x) - u(x - \delta)}{u(x)}, \frac{u(x)}{u(1 - \delta)}\}, \tag{5}$$

$$r_B^\delta(x) = \min\{a, b\}\,\min\{\frac{v(1 - x) - v(1 - x - \delta)}{v(1 - x)}, \frac{v(1 - x)}{v(1 - \delta)}\}, \text{ and} \tag{6}$$

$$r^\delta(x) = \min\{r_A^\delta(x), r_B^\delta(x)\}. \tag{7}$$

With respect to Part (i), only the second equation has changed. The reason is that Lk individuals draw their sample from their own population's past play if k is even, while they draw their sample from their rival population's past play if k is odd. Since $\hat{k}_A \geq 2$, both types of individuals are present in population A. Now it is not only the individuals from population B who may respond to mistakes by population A; individuals from population A themselves may also do so. For instance, an $L2$ individual from population A takes a sample from their own population, estimates from this sample the possible choices by the population B individual, and subsequently responds to that. This opens up the possibility of the $L2$ individual responding to the mistakes made by individuals in the own population. In case their sample is smaller (i.e., $a < b$) they may estimate a shift in population B whereas none of the individuals in population B would already consider a shift from the convention.

Note that for $a > b$, Equation (6) is identical to Equation (2), and we obtain the same

outcome as in Part (i). Next, consider the case $a \leq b$, for which

$$\frac{r^\delta(x)}{\delta} = \min\{a\,\frac{u(x)-u(x-\delta)}{\delta\,u(x)}, a\,\frac{u(x)}{\delta\,u(1-\delta)}, a\,\frac{v(1-x)-v(1-x-\delta)}{\delta\,v(1-x)}, a\,\frac{v(1-x)}{\delta\,v(1-\delta)}\}. \tag{8}$$

Now, $\lim_{\delta\downarrow 0}\frac{r^\delta(x)}{\delta}$ simplifies to $a\,\min\{\frac{u'(x)}{u(x)}, \frac{v'(1-x)}{v(1-x)}\}$, which is maximized at the (unique) solution to the equation $\frac{u'(x)}{u(x)} = \frac{v'(1-x)}{v(1-x)}$, that precisely characterizes the standard Nash bargaining solution.

Part (iii). Suppose we have sophisticated best responders in both populations. Now, there exists an individual in each population who samples the past actions of the rival population and there exists an individual in each population who samples the past action of its own population. The minimum number of mistakes needed to move away from the efficient convention $(x, 1-x)$ corresponds to the smallest integer greater than or equal to $m\,r^\delta(x)$, where

$$r_A^\delta(x) = \min\{a,b\}\,\min\{\frac{u(x)-u(x-\delta)}{u(x)}, \frac{u(x)}{u(1-\delta)}\}, \tag{9}$$

$$r_B^\delta(x) = \min\{a,b\}\,\min\{\frac{v(1-x)-v(1-x-\delta)}{v(1-x)}, \frac{v(1-x)}{v(1-\delta)}\}, \text{ and} \tag{10}$$

$$r^\delta(x) = \min\{r_A^\delta(x), r_B^\delta(x)\}. \tag{11}$$

Suppose, without loss of generality that $a \leq b$. Then,

$$\frac{r^\delta(x)}{\delta} = a\,\min\{\frac{u(x)-u(x-\delta)}{\delta\,u(x)}, \frac{u(x)}{\delta\,u(1-\delta)}, \frac{v(1-x)-v(1-x-\delta)}{\delta\,v(1-x)}, \frac{v(1-x)}{\delta\,v(1-\delta)}\}, \tag{12}$$

such that $\lim_{\delta\downarrow 0}\frac{r^\delta(x)}{\delta}$ simplifies to $a\,\min\{\frac{u'(x)}{u(x)}, \frac{v'(1-x)}{v(1-x)}\}$. From Part (ii) we know that this expression is maximized at the standard Nash bargaining solution. ∎

Young (1993b) shows, in a two-population world with only $L1$ individuals, that the bargaining power of a population is determined by its sample size, thereby implying that populations with smaller sample sizes are in a disadvantageous position. Parts (ii) and (iii) of Proposition 2 show that the presence of sophisticated individuals, levels out the disadvantage caused by drawing smaller samples. In particular, when both populations comprise of higher cognitive types (i.e., Lk individuals with $k \geq 2$), any disadvantage borne out of unequal sample size is "corrected" for and both populations end up with equal bargaining power.

4 General class of bimatrix games

In this section, we examine adaptive play between populations containing individuals of higher cognitive level in bimatrix games. We use the stochastically stable set of Young's adaptive play model (i.e., with $L1$ individuals only) as a point of reference and denote this set by $\bar{\Omega}^{a,b}$ for population A and B with sample sizes a and b respectively.

The next proposition deals with the case where only one of the populations (A) comprises of higher cognitive types. The situation where one population contains a share of $L2$

individuals and both populations have an equal sample size has been captured in Matros (2003).

Proposition 3. *Suppose that the highest level of cognition in population A is Lk with $k \geq 2$, while population B only contains L1 individuals. The stochastically stable set is described by $\bar{\Omega}^{a,b}$ if $a \geq b$ and by $\bar{\Omega}^{a,a}$ otherwise.*

Proof. From Proposition 1, we know that curb sets are invariant to cognitive hierarchies and that the unperturbed adaptive play process converges to a curb set.

Let ℓ be the minimum number of experimentations or mistakes required to transit from one recurrent class to another when there are only L1 individuals in either population. Moreover, let ℓ' be the minimum number of experimentations or mistakes required to transit from one recurrent class to another when there are only L1 individuals in population B while population A contains more sophisticated individuals (i.e., $\hat{k}_A \geq 2$).

The set of stochastically stable states might change if the minimum number of experimentations needed to transit from one recurrent class to another changes. Notice that if it is possible to get from one recurrent class to another in ℓ experimentations when $\hat{k}_A = 1$ and $\hat{k}_B = 1$, then it is certainly possible as well with ℓ experimentations when $\hat{k}_A \geq 2$ and $\hat{k}_B = 1$, since with positive probability only L1 individuals are chosen in population A. Therefore $\ell' \leq \ell$. The converse holds as well, i.e. $\ell \leq \ell'$. To see this, note that Lk individuals, with $k \geq 2$ odd, would play a strategy outside the prevailing minimal curb set only if there are enough number of mistakes in the rival population's past play. Thus, the minimum number of mistakes required to induce a best response outside the prevailing minimal curb set does not change. As a result, there are no transitions that take place in the presence of Lk individuals, with $k \geq 2$ odd, that would not take place with L1 individuals. Lk individuals, with $k \geq 2$ even, sample from their own population's past play and it would take a certain number of mistakes in their own population's past strategies for these individuals to play a best response not contained in the prevailing minimal curb set.

When sample sizes are equal (i.e., $a = b$), the same number of mistakes are sufficient to induce L1 individuals in population B to play a different strategy. Hence, the same transition can be established in a process with only L1 individuals. When population A has a larger sample size (i.e., $a > b$), L1 individuals in population B would in fact, owing to the smaller sample size, respond to a lesser number of mistakes in population A. The minimum number of experimentations for the transition involve L1 individuals only. Thus, the possibility of a similar transition also exists with only L1 individuals in either population. So, the set of stochastically stable states is $\bar{\Omega}^{a,b}$ if $a \geq b$.

Now, suppose population A has a smaller sample size (i.e., $a < b$). Then, Lk individuals in population A, with $k \geq 2$ even, react to the mistakes that appear in their own population earlier than the individuals in population B. The L1 individuals in population B would

13

have reacted to the same mistakes if they had a sample size equal to a. As a result, the same transition would be effected with only $L1$ individuals in both populations, where both populations are endowed with sample size a. Hence, the set of stochastically stable states is $\bar{\Omega}^{a,a}$ if $a < b$. ∎

Proposition 3 implies that the presence of individuals with higher cognition only influences the long run outcome when the population that they belong to is endowed with a smaller sample size. In such a case, the effect of these individuals with higher cognition is to "decrease" the sample size of the rival population to that of the own population. When both populations comprise of higher cognitive types, the net effect is to "equalize" the sample size of both populations. This is precisely what the next proposition states.

Proposition 4. *Suppose both populations have individuals with a level of cognition of at least 2. The stochastically stable set is $\bar{\Omega}^{a,a} = \bar{\Omega}^{b,b}$.*

Thus the above two propositions generalize the result obtained for the evolutionary Nash bargaining model.

5 Extensions

In this section we present two extensions of the model considered. The first considers an alternative specification of level-k reasoning, where we assume an individual of level k to believe the opponent may be of any lower level with positive probability. In the second we assume that individuals of higher cognition (falsely) use their own utility function when forming a believe on the rival population's play.

5.1 Alternative specification of level-k reasoning

In this first extension, we replace the assumption that Lk individuals believe that the opponent is of level $k-1$ and that each population comprises a positive mass of any lower type than the maximum. Instead, we assume now that an Lk individual believes that the opponent is of at most level $k-1$, with positive mass on all cognitive levels up to $k-1$.[9] Let us denote this individual by L^*k.

The L^*k individual recognizes the fact that the rival could be of any level between and including 1 and $k-1$ with positive probability and considers all best responses to any of these possible opponents. This alternative specification of level-k reasoning does not affect the results stated in Propositions 3 and 4 of the preceding section. To see this, it suffices to realize that since an L^*k individual places a positive weight on all lower cognitive types, she behaves like an Lk (when a best response corresponding to the belief that the opponent is an $L(k-1)$ individual is chosen) or an $L(k-1)$ individual (when a best response corresponding to the

---

[9]Note that we do not explicitly require all these types actually to be contained in the rival population.

14

belief that the opponent is an $L(k-2)$ individual is chosen), each with positive probability. Hence, the minimum number of experimentations to transit from one minimal curb set to another does not change in comparison to the model with Lk individuals.

Recall that we do not require L^*k individuals to assign precise probabilities on the likelihood to meet a rival of particular (lower) level of cognition, as is done in Stahl's (1993) model of sophistication and in Camerer et al.'s (2004) model of cognitive hierarchies. The latter assumes individuals to perfectly know the fraction of lower cognitive types in the population, such that the beliefs of individuals with higher level of cognition are closer to the actual proportions in the population. It is straightforward to see that these restrictions on the beliefs, do not affect the recurrent classes (which are the minimal curb sets) and the stochastically stable set as long as the perceived probability to meet a type of any lower cognitive level is strictly positive. For Camerer et al.'s (2004) model of cognitive hierarchies it would require all lower cognitive types to be actually present in the rival population, thereby only allowing for the cases for which $|\hat{k}_A - \hat{k}_B| \leq 1$.

5.2 Knowledge of utilities

One source of dissatisfaction with the model in the previous sections might be that individuals with level of cognition higher than or equal to 2 know the utility function of the individuals in the other population. In this subsection, we relax this assumption for the class of games G', where the strategy sets and the ordinal preferences are identical for all individuals in either population.[10] We assume that individuals only know their own (population's) utility function and not that of their possible rivals. This implies that though the ordinal preferences are the same across the two populations, individuals in one population are not aware of the cardinal utilities of the individuals in the other population. We denote the resulting individual with level-k depth of reasoning by $L'k$. Note, though, that we retain the assumption that if a population contains an $L'k$ individual ($k \geq 2$), then it also contains an $L'(k-1)$ individual. Let $\hat{k}'_A$ and $\hat{k}'_B$ denote the highest cognitive level in population A and population B respectively.

Like Lk individuals, $L'k$ individuals draw samples from the strategies of the rival population when k is odd and of the own population when k is even. So, $L'1$ individuals behave identically to $L1$ individuals: they best respond to the empirical distribution of strategies in the sample drawn from the opponent population's past play. $L'2$ individuals best respond to their estimate of their rival's behavior. Assuming that their rival is an $L'1$ individual, they sample from their own population's past play to assess the strategies the rival may be using. However, for this assessment they use their own utility function rather than that of the rival.[11]

---

[10]Examples outside this class for which the propositions below do not hold are easily constructed.

[11]Possible explanations of such a systematic behavioral trait include the 'false consensus effect' (Ross et al., 1977) and self-projection (Buckner and Carroll, 2007).

The following proposition claims that the resulting unperturbed adaptive play process is invariant to this modification of the level-k model.

Proposition 5. *In the class of games G', if the history m is sufficiently large and the sampling is sufficiently incomplete (i.e., a and b are sufficiently small), then the process of adaptive play converges to a minimal curb set with probability one.*

Proof. The proof is a trivial modification of the proof of Proposition 2. ∎

For further selection among the recurrent classes (i.e., the minimal curb sets) of the unperturbed process, we again adopt the notion of stochastic stability by studying the support set of the limit of the stationary distribution of regular perturbations of the process when the perturbation vanishes. Despite the invariance of the recurrent classes with respect to the current alternative specification of level-k behavior, the stochastically stable set may change. The next proposition illustrates how the stochastically stable state of the evolutionary Nash bargaining game is changed.

Proposition 6. *(i) Suppose that population A contains a fraction of $L'k$ individuals for $k \geq 2$, while population B is comprised by $L'1$ individuals only. The stochastically stable state is either the generalized Nash bargaining solution with the bargaining power of a population equal to its own sample proportion, or the 50–50 division (equal split).*

(ii) Suppose that both populations comprise of a fraction of $L'2$ individuals. The stochastically stable state is either the generalized Nash bargaining solution with the bargaining power of a population equal to its own sample proportion, a modified generalized Nash bargaining solution with the bargaining power of a population equal to the other population's sample proportion, or the 50–50 division.

Proof. By Proposition 5, play settles in a minimal curb set, which corresponds to an efficient convention. For stochastic stability we need to examine the relative ease or difficulty of transiting from an efficient convention.

Part (i). An efficient convention $(x, 1-x)$ can be disrupted by mistakes by population A or by mistakes by population B.

First, consider mistakes by population A. One possibility is that these mistakes affect the sample of $L'1$ individuals in population B. For these individuals to change their demand the minimum number of mistakes required equals $r_{BA}^\delta(x) = b \min\{\frac{v(1-x)-v(1-x-\delta)}{v(1-x)}, \frac{v(1-x)}{v(1-\delta)}\}$. The other possibility is that these mistakes are sampled by $L'k$ individuals in population A with k even. For these individuals to change their demand the minimum number of mistakes required equals $r_{AA}^\delta(x) = a \min\{\frac{u(1-x)-u(1-x-\delta)}{u(1-x)}, \frac{u(1-x)}{u(1-\delta)}\}$. After all, these individuals would demand something apart from x if they believe that an $L'(k-1)$ individual in population B is going to demand something else than $1-x$. The latter possibility occurs if the $L'(k-1)$ individual in population B believes that an $L'(k-2)$ individual in population A is going to demand

16

something different than x, and so on. Ultimately, this can be traced to the belief that an $L'1$ individual in population B changes the demand. If this does not happen, then none of the higher cognitive individuals in any population believe that the action of the rival is going to change. Individuals in population A form this belief when they sample a sufficient number of mistakes in their own population. Since we are looking at the minimum number of mistakes to move out of a convention, it suffices to examine the case when the mistakes in their own population are such that the mistaken demands are just more than x (i.e., $x + \delta$) or when they ask for the least amount (i.e., δ). Suppose that ℓ mistakes of the first type have been made. For an individual A to consider it viable that population B may change (to $1 - x - \delta$), ℓ should be such that $u(1 - x - \delta) \geq \frac{a\,m - \ell}{a\,m} u(1 - x)$. This gives the least proportion of mistakes necessary to be $\frac{\ell}{m} = a\,\frac{u(1-x) - u(1-x-\delta)}{u(1-x)}$. The second term can be derived by considering the second type of mistakes.

Next, consider mistakes by population B. These mistakes can only affect the sample of $L'k$ individuals in population A with k odd. For these individuals to change their demand the minimum number of mistakes is required to be $r_{AB}^{\delta}(x) = a\,\min\{\frac{u(x) - u(x - \delta)}{u(x)}, \frac{u(x)}{u(1-\delta)}\}$.

The minimum resistance of the efficient convention $(x, 1 - x)$ is thus given by $r^{\delta}(x) = \min\{r_{BA}^{\delta}(x), r_{AA}^{\delta}(x), r_{AB}^{\delta}(x)\}$. It is easily seen (by using the same argument as in the proof of Proposition 2) that $\lim_{\delta \downarrow 0} \frac{r_{\delta}(x)}{\delta} = \min\{a\,\frac{u'(x)}{u(x)}, a\,\frac{u'(1-x)}{u(1-x)}, b\,\frac{v'(1-x)}{v(1-x)}\}$. The first of these terms is decreasing in x while the latter two are increasing in x. The stochastically stable efficient allocation $(\bar{x}, 1 - \bar{x})$ is at the (unique) $\bar{x}$ that maximizes $\frac{r_{\delta}(x)}{\delta}$, which is at the unique solution to $a\,\frac{u'(x)}{u(x)} = \min\{a\,\frac{u'(1-x)}{u(1-x)}, b\,\frac{v'(1-x)}{v(1-x)}\}$. Depending on which of the two terms in the right-hand-side is lowest at $\bar{x}$, we either find the 50–50 division, or the generalized Nash bargaining solution with the power of the population equal to its own sample size.

Part (ii). An efficient convention $(x, 1 - x)$ can be disrupted by mistakes by population A or by mistakes by population B.

First, consider mistakes by population A. One possibility is that these mistakes affect the sample of $L'k$ individuals in population B with k odd. For these individuals to change their demand the minimum number of mistakes required equals $r_{BA}^{\delta}(x) = b\,\min\{\frac{v(1-x) - v(1-x-\delta)}{v(1-x)}, \frac{v(1-x)}{v(1-\delta)}\}$. The other possibility is that these mistakes are sampled by $L'k$ individuals in population A with k even. For these individuals to change their demand the minimum number of mistakes required equals $r_{AA}^{\delta}(x) = a\,\min\{\frac{u(1-x) - u(1-x-\delta)}{u(1-x)}, \frac{u(1-x)}{u(1-\delta)}\}$.

Next, consider mistakes by population B. One possibility is that these mistakes affect the sample of $L'k$ individuals in population A with k odd. For these individuals to change their demand the minimum number of mistakes required equals $r_{AB}^{\delta}(x) = a\,\min\{\frac{u(x) - u(x-\delta)}{u(x)}, \frac{u(x)}{u(1-\delta)}\}$. The other possibility is that these mistakes are sampled by $L'k$ individuals in population B with k even. For these individuals to change their demand the minimum number of mistakes required equals $r_{BB}^{\delta}(x) = b\,\min\{\frac{v(x) - v(x-\delta)}{v(x)}, \frac{v(x)}{v(1-\delta)}\}$.

The minimum resistance of the efficient convention $(x, 1 - x)$ is thus given by $r^{\delta}(x) =$

$\min\{r_{BA}^{\delta}(x), r_{AA}^{\delta}(x), r_{AB}^{\delta}(x), r_{BB}^{\delta}(x)\}$. It is easily seen (by using the same argument as in the proof of Proposition 2) that $\lim_{\delta \downarrow 0} \frac{r_{\delta}(x)}{\delta} = \min\{a\frac{u'(x)}{u(x)}, a\frac{u'(1-x)}{u(1-x)}, b\frac{v'(x)}{v(x)}, b\frac{v'(1-x)}{v(1-x)}\}$. The first and third of these terms are decreasing in x while the other two are increasing in x. The stochastically stable efficient allocation $(\bar{x}, 1 - \bar{x})$ is at the (unique) $\bar{x}$ that maximizes $\frac{r_{\delta}(x)}{\delta}$, which is at the unique solution to $\min\{a\frac{u'(x)}{u(x)}, b\frac{v'(x)}{v(x)}\} = \min\{a\frac{u'(1-x)}{u(1-x)}, b\frac{v'(1-x)}{v(1-x)}\}$. Depending on which of the two terms in the two sides of the equation are lowest at $\bar{x}$, we either find the 50–50 division, the generalized Nash bargaining solution with the power of the population equal to its own sample size, or the generalized Nash bargaining solution with the power of the population equal to the other population's sample size. ∎

Proposition 6 shows that this alternative specification of level-k behavior has a severe qualitative impact on the long run outcome. The $L'k$ specification allows for the equal split and an inverted version of the generalized Nash bargaining solution to be feasible long run outcomes. This latter is not a long run outcome in the Lk specification, while the former was found only when the cardinal utilities of the populations were identical up to an affine transformation.

6   Concluding remarks

In this paper we investigated the effect of sophisticated agents on adaptive play. Apart from the more usual way of defining sophisticated agents as higher order best responders (Lk individuals), we introduced an alternative notion that does not assume (higher order best responding) individuals to be aware of the utility function of the individuals in the rival population ($L'k$ individuals). We find that the unperturbed adaptive play process, with either notion of sophistication, converges to a minimal curb set. The intuition for this result is simple. With only $L1 = L'1$ individuals, best response adaptive play converges to a minimal curb set in a finite number of steps. So, since with positive probability only simple best responders are chosen for any finite number of steps, the process converges to a minimal curb set. By the construction and definition of a minimal curb set, higher order best responses are contained in it. Hence, any strategy chosen by a more sophisticated best responder would be in the minimal curb set.

In addition to the focus on the recurrent classes, we also considered the effect of these more sophisticated best responders on the stochastically stable outcomes relative to Young's setting with only simple best responders. We find that higher level of sophistication has no effect when the population containing these has the same or higher sample size. However, sophistication has a differential effect when the population holding these individuals has a strictly smaller sample size. The effect is as if to reduce the sample size of the rival population to that of the own population. This effect is already obtained by the presence of a fraction of $L2$ individuals and the addition of even higher cognitive types have no further effect. Again, the intuition is simple. Higher sophistication has an influence only when it results in a formation of a belief

that the rival is going to change her behavior, and the sophisticated best responders actually changing their own behavior on basis of that belief. An individual needs to be at least of level 2 to form such a belief. Only when the sophisticated best responders have a smaller sample size the belief is "unfounded", which is why that is the only situation in which they have a differential effect on the stochastically stable outcomes. When both populations have $L2$ individuals, then it is as if we are dealing with the basic adaptive play model of Young with both populations having equal sample size. The presence of $L2$ individuals on both sides, level out any differences in sample size. And, again, more sophistication beyond $L2$ types has no additional effect – as mentioned earlier, all higher beliefs formed can be traced to the behaviour of an $L1$ type individual. With $L1$ and $L2$ individuals, the mistakes in either populations' play that may give rise to different beliefs are accounted for. Hence, even higher cognition has no effect.

We hazard a guess and suspect that these properties are transferable to a setting with multiple populations. First, convergence to a minimal curb set is clearly obtained for any number of populations and any distribution of cognitive types across populations. Moreover, regarding stability against experimentation, still all plays of strategies outside a prevailing curb set can be traced to either a level-1 individual in some population to change play with positive probability (based on a sufficient number of mistakes), or there should be a belief by some higher order type that this is about to happen.

Throughout the paper, we have restricted attention to the best reply dynamic and have ignored the weakly better reply dynamic that is less demanding with respect to the individuals' rationality. One attractive feature of considering weakly better replies is that under a wide class of evolutionary dynamics, a set is asymptotically stable if and only if it is closed under the better reply correspondence (Ritzberger and Weibull, 1995, Thm. 1). With level-1 individuals, play converges to a set that is closed and fixed under the better reply correspondence. If we incorporate the element of higher cognition, then for the very same reasons as in Proposition 1, this result is unaffected. However, it is of course possible that the set of stochastically stable states varies.

Having identified the effect that higher cognition has on adaptive play, the next question is their qualitative impact: Do more sophisticated individuals leave their population better off? Thereto, suppose that population A consist of $L1$ and $L2$ individuals, while population B has $L1$ individuals only. Suppose further that the sample size of population B is greater than that of population A ($a < b$), so that by Proposition 4, the stochastically stable states are the same as in a process where both populations have simple best responders only and the sample sizes are equal. We argue that depending on the payoffs of the game, this outcome may be better or worse for population A – the population with the clever agents. That population A may benefit from the presence of clever agents was seen in the Nash bargaining model. However, this is not a universal phenomenon. To this end, we draw attention to an example

from Young (1998, Ch. 5) that illustrates that having less information (i.e., a smaller sample size and therefore higher sensitivity to mistakes) may be an advantage. Consider the game in Figure 1 and assume that the row population has a sample size of s while the column population has a sample size of s', with $\frac{9}{19}\,s < \frac{7}{17}\,s'$. When both populations have simple best

$$
\begin{array}{c|cc}
 & L & R \\
\hline
U & 10,7 & 0,0 \\
D & 0,0 & 9,10 \\
\end{array}
$$

Figure 1: Example

responders only, then the outcome with payoffs $(10,7)$ is stochastically stable – i.e., the row population is better off. However, if the row population contains a fraction of sophisticated best responders (i.e., $L2$ individuals), then the outcome with payoffs $(9,10)$ is stochastically stable. Hence, the row population suffers from the presence of higher cognitive individuals in their population.

An interesting qualitative impact of sophistication comes from Matros (2003): Sophistication might influence the distribution of strategies inside the stochastically stable minimal curb set(s). This is illustrated by Matros (2003) in the game of Matching Pennies, where one of the populations had a fraction of $L2$ individuals. In the long run outcome, these $L2$ individuals have a positive expected payoff while the $L1$ individuals in both populations have a negative expected payoff. We suspect that even though sophistication beyond $L2$ has no effect on the stochastically stable minimal curb set(s), similar features may be observed. There may also be differences in the speed of convergence. The finding of Matros (2003) raises another interesting issue: Will more sophisticated types end up with a payoff at least as large as that of less sophisticated types? We suspect that this will not generally be the case in a model where there are at least $L2$ types on both populations as the expected payoffs will depend heavily on the precise proportion of each type in each population.

References

[1] Kaushik Basu and Jörgen Weibull (1991). Strategy subsets closed under rational behavior. Economic Letters 36 (2): 141-146.

[2] Kenneth Binmore (1987). Modeling rational players I. Economics and Philosophy 3: 179-214.

[3] Kenneth Binmore (1988). Modeling rational players II. Economics and Philosophy 4: 9-55.

[4] Randy L Buckner and Daniel C Carroll (2007). Self-projection and the brain. Trends in Cognitive Sciences 11 (2): 49-57.

[5] Colin F Camerer, Tech-Hua Ho and Juin-Kuan Chong (2004). A cognitive hierarchy model of games. Quarterly Journal of Economics 119 (3): 861-898.

[6] Giorgio Coricelli and Rosemarie Nagel (2009). Neural correlates of depth of strategic reasoning in medial prefrontal cortex. Proceedings of the National Academy of Sciences 106 (23): 9163-9168.

[7] Vincent P Crawford and Nagore Iriberri (2007). Fatal attraction: Salience, naivete, and sophistication in experimental "hide-and-seek" games. American Economic Review 97 (5): 1731-1750.

[8] Sjaak Hurkens (1995) Learning by forgetful players. Games and Economic Behaviour. 11 (2): 304-329.

[9] Alexander Matros (2003). Clever agents in adaptive learning. Journal of Economic Theory 111 (1): 110-124

[10] Erik Mohlin (forthcoming). Evolution of theories of mind. Games and Economic Behavior, forthcoming.

[11] Rosemarie Nagel (1995). Unraveling in guessing games: An experimental study. American Economic Review 85 (5): 1313-1326.

[12] Klaus Ritzberger and Jörgen Weibull (1995). Evolutionary selection in normal-form games. Econometrica 63 (6): 1371-1399.

[13] Lee Ross, David Greene and Pamela House. (1977). The 'false consensus effect': An egocentric bias in social perception and attribution processes. Journal of Experimental Social Psychology 13 (3): 279-301.

[14] María Sáez-Martí and Jörgen Weibull (1999). Clever agents in Young's evolutionary bargaining model. Journal of Economic Theory 86 (2): 268-279.

[15] Dale O Stahl (1993). Evolution of $smart_n$ players. Games and Economic Behaviour 5 (4): 604-617.

[16] Dale O Stahl and Paul W Wilson (1994). Experimental evidence on players' models of other players. Journal of Economic Behavior & Organization 25 (3): 309-327.

[17] Dale O Stahl and Paul W Wilson (1995). On players' models of other players: Theory and experimental evidence. Games and Economic Behaviour 10 (1): 218-254.

[18] Joseph T-Y Wang, Michael Spezio and Colin F Camerer (2010). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. American Economic Review 100 (3): 984-1007.

[19] H Peyton Young (1993a). The evolution of conventions. Econometrica 61 (1): 57-84.

[20] H Peyton Young (1993b). An evolutionary model of bargaining. Journal of Economic Theory 59 (1): 145-168.

[21] H Peyton Young (1998). Individual Strategy and Social Conventions. Princeton University Press.