Deep and shallow thinking in the long run*

Heinrich Nax† and Jonathan Newton‡

March 25, 2021

Abstract

Humans differ in their strategic reasoning abilities and in beliefs about others' strategic reasoning abilities. Studying such cognitive hierarchies has produced new insights regarding equilibrium analysis in economics. This paper investigates the effect of cognitive hierarchies on long run behavior. Despite short run behavior being highly sensitive to variation in strategic reasoning abilities, this variation is not replicated in the long run. In particular, when generalized risk dominant strategy profiles exist, they emerge in the long run independently of the strategic reasoning abilities of players. These abilities may be arbitrarily low or high, heterogeneous across players and evolve over time.

Keywords: bounded rationality, level-*k* thinking, evolution.

JEL Codes: C73, D81, D90.

1. Introduction

> *"Coordination, when it occurs, is an almost accidental (though statistically predictable) by-product of non-equilibrium thinking"*
>
> – Vincent Crawford (2007)

There is evidence that humans sometimes reason iteratively to predict the behavior of others and that the depth of such reasoning can vary according to person and situation (Crawford, 2019). Apart from some notable exceptions (e.g. Sáez-Martı and Weibull, 1999; Myatt

*The authors gratefully acknowledge the comments of Xiaohong Chen, Vince Crawford, Ryota Iijima, Rosemarie Nagel, María Sáez Martí and Larry Samuelson. Nax's work was supported by a Swiss National Science Foundation Eccellenza Grant. Newton's work was supported by a KAKENHI Grant-in-Aid for Research Activity Start-Up funded by the Japan Society for the Promotion of Science (Grant 19K20882).

†ETH Zurich, 8092 Zurich, Switzerland, and University of Zurich, 8050 Zurich, Switzerland. e-mail: `heinrich.nax@uzh.ch`; website: `nax.science`.

‡Institute of Economic Research, University of Kyoto, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan. e-mail: `newton@kier.kyoto-u.ac.jp`; website: `jonathannewton.net`.

Level of rationality	Risk dominance with two strategies and uniform interaction	Generalized risk dominance
$k = 1$	Blume (2003)	Peski (2010)
$2 \leq k < \infty$	$*$	$*$
$k = \infty$	Myatt and Wallace (2003)	$*$

Figure 1: **Summary of literature.** Blume (2003) and Peski (2010) consider myopic ($k = 1$) players that follow the general class of processes that we consider in Section 6. Myatt and Wallace (2003) consider sophisticated ($k = \infty$) players who can iteratively reason to a Nash equilibrium every period. The current paper (denoted $*$) extends these results to any k in the generalized setting and, a fortiori, to the two strategy uniform interaction setting. Other work, discussed later, considers two player games under sample based processes (Sáez-Martı and Weibull, 1999; Matros, 2003; Khan and Peeters, 2014).

and Wallace, 2003), the literature on long run outcomes in games played in populations usually abstracts from such considerations. An open question has been whether the best known result in this literature, the emergence of (generalized) risk dominant Nash equilibria (see Peski, 2010) under broad classes of best response dynamics, is robust to such iterative reasoning. We answer this question in the affirmative. Even though short run behavior can be dramatically affected by such reasoning and convergence to Nash equilibrium may fail to occur in the short run, long run predictions are robust to all levels of reasoning by players. Moreover, these levels may be heterogeneous and may even be random, in which case they can be correlated across players.

Let us describe our model in more detail. Every period, given the current strategy profile, each player formulates a conjecture about the behavior of the other players to which he will usually, but not always, best respond. A player of level $k = 1$ will conjecture that other players remain at the current strategy profile. Higher levels of k are defined iteratively. A player of level k will conjecture that all other players are of level $k-1$. Level $k = \infty$ involves reasoning to a Nash equilibrium strategy every period. We consider the long run behavior of this process. The long run stability of risk dominance has been known in progressively stronger forms since Young (1993a); Kandori et al. (1993). The strongest known results for various levels of k in specific and general settings are shown in Figure 1, with the remaining entries in the table being contributions of the current paper. In addition, our results span the table in the sense that if, as empirical work suggests (Stahl and Wilson, 1994; Nagel, 1995, onwards) and is theoretically plausible (Stahl, 1993), players have different values of k, or if levels of k are determined randomly from period to period, then the result still holds.[1]

[1]Notably, in addition to epistemic considerations (Bacharach, 1992), the foundational literature on level k thinking explicitly studies evolutionary forces (Stahl, 1993) and behavioral dynamics (Nagel, 1995; Selten,

For expositional reasons, we first state our main result for a level k version of logit choice (Theorem 1). Later, we expand our analysis to consider a broad class of dynamics in which the probability of playing a non-best response is weakly decreasing in the payoff loss that results, together with a broad class of conjectures that players can make about the behavior of other players (Theorem 3). Lastly, we consider randomness in players' conjectures (Theorem 4). To prove these results, we use recent advances in the study of asymmetric dynamics. We break down the dynamic process of strategy updating into sub-processes for each player and level of rationality and show that these sub-processes satisfy a certain type of asymmetry towards the generalized risk dominant strategy profiles. We then combine these processes using the methods of Newton (2020, 2019) to obtain an aggregate process that is also asymmetric towards such profiles. This asymmetry implies that generalized risk dominant strategy profiles are those that will be observed most often in the long run (Peski, 2010).[2]

Applying our results to technology adoption in networked populations, we show that the strategy profile at which every player plays a risk dominant technology remains the most stable outcome for arbitrary levels of rationality (Proposition 1). We then consider the possibility of entirely random ($k = 0$) players and show how their presence can cause multiplicity in the number of long run stable outcomes (Theorem 2) in a way that depends subtly on the levels of more rational ($k \geq 1$) players (Proposition 2). Applying our results to the second price auction, we show that the strategy profile at which every player bids his valuation emerges in the long run (Proposition 3) for a broad range of conjectures that players can form about how other players choose their bids. This result is proven by showing that an equilibrium in weakly dominant strategies is always generalized risk dominant (Theorem 5).

The paper is organized as follows. Section 2 discusses related literature. Section 3 describes level k logit dynamics and Section 4 gives the first iteration of our main result. Section 5 applies this result to technology adoption on networks. Section 6 gives our main result under a broad class of dynamics and a broad class of conjectures that players can make about the behavior of other players. Section 7 applies this result to second price auctions. All proofs are relegated to the appendix.

1991). The current study and those cited in Figure 1 follow in this tradition by studying long run behavior under such models.

[2] Theorem 1 on level k conjectures and logit choice is, in fact, a special case of the subsequent theorems on general conjectures and general dynamics. However, given the accessibility and widespread use of level k models and logit choice, we present these results prior to the more abstract and general presentation.

2. Related literature

Economic experiments on static, sequential and repeated games have produced evidence to support both level k modeling assumptions and the possibility of heterogeneity in k across players. See Crawford (2019) for a recent review of this literature. Moreover, there is evidence that short run convergence to equilibrium may depend on cognitive ability (Gill and Prowse, 2016; Proto et al., 2019). However, there is also experimental evidence of a tendency for subjects to reason at progressively higher levels over time (Nagel, 1995; Duffy and Nagel, 1997). Indeed, changes in levels of k have been fitted to models of reinforcement learning Stahl (1999, 2000) and Bayesian updating (Ho and Su, 2013). In contrast, the main results of the current paper concern long run behavior after levels of k have either reached a steady state or persist above a given level. In such situations, our results predict a tendency towards risk dominant equilibria, as has been observed in several experimental settings (Van Huyck et al., 1990; Battalio et al., 2001; Heinemann et al., 2004; Cabrales et al., 2007).

There also exists a theoretical literature on the persistence of different levels of k in populations. The overall conclusion is that because "being right is just as good as being smart" (Stahl, 1993), heterogeneous levels of rationality can persist when players with those levels play similar strategies (see also Mohlin, 2012). Other work discusses explicit weaknesses of iterated reasoning. Stennek (2000) shows how iterated deletion of strictly dominated strategies can lead to fitness losses unless probability weight redistributed from a dominated strategy when it is deleted is only transferred to those strategies that dominate it. Geanakoplos and Gray (1991) explain how errors in assessing the value of future continuation games can lead to suboptimal play in the present. A striking way to be unable to assess the value of the future in such settings would be if a player did not use information about other player's payoffs when making decisions. As we might expect, such players will usually be eliminated from the population (Robalino and Robson, 2016). Finally, heterogeneity in other traits may interact with iterated reasoning. For example, Heller (2015) shows how being able to know far in advance when a series of repeated prisoner's dilemmas will end can be evolutionarily selected against, as when two such players are paired, it becomes impossible to sustain cooperation for a considerable number of periods before the end of the game. Similarly, collaborative decision making may lead to cooperation in prisoner's dilemmas in the absence of further reasoning, but will fail to do so if subsequent reasoning lead to defection (Newton, 2017; Rusch, 2019).

The dynamic processes that we adapt for level k and broader conjectures about opponents' behavior are common in the evolutionary literature. Theorem 1 concerns logit dynamics (Blume, 1993). For a detailed discussion of logit dynamics, the reader is directed to Alós-Ferrer and Netzer (2010). The general idea is that the probability of playing a non-best response action decreases log-linearly in the payoff loss from playing that action relative to the payoff from playing a best response. If log-linearity is dropped, we obtain a much larger class of dynamics. Theorems 3 and 4 concern this class, which is close to the classes of *skew-symmetric* rules (Blume, 2003) and *payoff-based* rules (Peski, 2010). Recent experimental studies designed to explicitly test the properties of non-best response behavior finds evidence in support of such dynamics (Mäs and Nax, 2016; Lim and Neary, 2016; Hwang et al., 2018). The survey of Newton (2018) covers recent work on trait evolution and dynamics, including many of the papers discussed above.

The dynamics considered in the current paper have the current strategy profile as the state variable. This has been popular in the literature following Kandori et al. (1993). A parallel literature considers a sample based process, adaptive play (Young, 1993a), according to which members of populations are drawn to play a game against members of other populations and respond to a sample of how the game has been played in the recent past. This adds a degree of complication to the model that has been leveraged to obtain results for level k thinking in two player games in which one player is drawn from each of two populations. From a benchmark in which every player has $k = 1$, the focus of research has been on conditions under which the presence of $k = 2$ players changes the implications of the benchmark model. Sáez-Martı and Weibull (1999) consider the bargaining model of Young (1993b) in the presence of $k = 2$ players, Matros (2003) considers generic two player games in the presence of $k = 2$ players, and Khan and Peeters (2014) consider generic two player games in the presence of players with any finite k. The general conclusion of this literature is that $k > 1$ makes a difference if and only if a low sample size for clever players in one population causes them to foresee a change in the behavior of the opposing population the next period that does not in fact happen. However, having acted to preempt the predicted change, the clever players put in motion a sequence of transitions that moves the process to another equilibrium.

Finally, we note that there is a literature that considers possible non-convergence of certain processes to mixed Nash equilibria (see Crawford, 1974) and the role that $k = 2$ players can have in overcoming this non-convergence (Selten, 1991; Conlisk, 1993b,a; Tang, 2001). In contrast, the dynamics of the current paper may fail to converge to Nash equilibrium in the short run due to the presence of $k = 2$ players. However, our main results regarding long run predictions turn out to be unaffected by whether or not short run convergence occurs.

3. Model

3.1 The game

Consider a normal form game $G = (N, \{S_i\}_{i\in N}, \{U_i\}_{i\in N})$. The set of players is N. Each player $i \in N$ has a finite strategy set S_i and a strategy for player i is denoted $s_i \in S_i$. The set of strategy profiles is $S = \times_{i\in N} S_i$ with generic element $s \in S$. Let s_{-i} denote s restricted to $N \setminus \{i\}$. The payoff to player i at strategy profile s is given by $U_i(s) = U_i(s_i, s_{-i})$. Assume that $U_i(s_i, s_{-i}) \neq U_i(s_i', s_{-i})$ for any s_i, s_i', s_{-i}, which holds generically in payoffs. Cases in which it does not hold will be considered later in Sections 6 and 7.

3.2 Level k best responses

We assume that each player $i \in N$ has a level of rationality given by an integer $k \geq 1$. Different players are allowed to have different values of k. We refer to a player with a given value of k as a *level k player*. A player's level will determine the conjecture he makes about the behavior of other players. In Section 6 we will allow players' levels to be random, correlated and changing over time. For now, we assume that any given player's level remains fixed and unchanging.

For a given strategy profile s, we denote the profile of best responses by

$$B^1(s) = \Big(B_i^1(s)\Big)_{i\in N}, \quad \text{where} \quad B_i^1(s) \in \arg\max_{s_i\in S_i} U_i\Big(s_i, s_{-i}\Big). \tag{1}$$

We will refer to $B^1(s)$ as the profile of *level* 1 best responses to s. Note that, by our genericity assumption on payoffs, best responses are uniquely determined. We also wish to consider best responses to best responses, best responses to these in turn, and so on. To do this, we recursively define *level k* best responses as

$$B^k(s) = B^1\Big(B^{k-1}(s)\Big) \quad \text{for} \quad k \in \mathbb{N},\ k \geq 2. \tag{2}$$

Note that the difference between different levels of best response lies in the conjectured strategy profiles to which a player best responds. These conjectures are based on the iteration of the best response correspondence. More general conjectures that are not based on iterated best response will be considered in Section 6.

If the best response correspondence converges, that is if there exists k such that $B^k(s) = B^{k-1}(s)$ for all $s \in S$, then we say that the game G is *Nash convergent*.

Remark 1. If G is Nash convergent, then for all $s \in S$, $B^\infty(s) := \lim_{k\to\infty} B^k(s)$ is a Nash

$$i \text{ is } \begin{cases} \text{active w.p. } q_i & \text{and plays } s_i' \text{ w.p.} \begin{cases} Pr^1(s_i'|s^t) & \text{if } i \text{ is level 1} \\ Pr^k(s_i'|s^t) & \text{if } i \text{ is level } k \geq 2 \end{cases} \\ \text{inactive w.p. } (1-q_i) & \text{and plays } s_i^{t+1} = s_i^t \end{cases}$$

$\xrightarrow{\quad}$ Period t $\quad$ $\xrightarrow{\qquad\qquad}$ Period $t+1$ $\quad$ $\xrightarrow{\quad}$ Period $t+2$

Figure 2: **Strategy updating.** Revision probabilities in period $t+1$ for different levels of rationality of player i.

equilibrium (Nash, 1950).

3.3 Level k logit choice

For our principal analysis we will adopt the logit choice rule. We do this as it is the best known and most commonly used stochastic behavioral rule. However, our results extend to a family of payoff dependent behavioral rules, which we shall discuss in Section 6. The standard logit choice rule (Cox, 1958; Block and Marschak, 1959) is as follows. Given a current strategy profile s^t, the probability of player i choosing strategy s_i' is given by

$$Pr^1(s_i'|s^t) = \frac{e^{\frac{1}{\eta}U_i\left(s_i', s_{-i}^t\right)}}{\sum_{s_i \in S_i} e^{\frac{1}{\eta}U_i\left(s_i, s_{-i}^t\right)}}, \quad \text{for some } \eta > 0. \tag{3}$$

This is a perturbed best response rule parameterized by $\eta > 0$. For small values of η, a player following this rule will usually play a level 1 best response. Hence, we refer to the rule as *level* 1 *logit choice*. However, sometimes the player will play a non-best response. For small values of η, such non-best responses are rare and the level 1 best response is played almost all of the time. Analogously, we define *level k logit choice* for $k \geq 2$,

$$Pr^k(s_i'|s^t) = \frac{e^{\frac{1}{\eta}U_i\left(s_i', B_{-i}^{k-1}(s^t)\right)}}{\sum_{s_i \in S_i} e^{\frac{1}{\eta}U_i\left(s_i, B_{-i}^{k-1}(s^t)\right)}}. \tag{4}$$

That is, for small values of η, a player following the level k logit choice rule will usually play a level k best response.

For $\eta = 0$, define level k logit choice probabilities as the limits of (3) and (4) as $\eta \to 0$. That is, a level k best response will be played with probability one.

The difference between the standard logit choice rule and the level k logit choice rule for $k \geq 2$ is the conjectured play of the opposing players. Specifically, standard logit probabilities for player i are calculated with respect to the conjecture that other players remain playing their current strategies, whereas level k logit probabilities are calculated with respect to the conjecture that other players play level $k-1$ best responses.[3]

It follows from (4) and the definition of Nash convergence that if the game G is Nash convergent and players are sufficiently rational, then level k logit choice under small η will usually conform to the play of Nash equilibrium strategies.[4]

Remark 2. If the game G is Nash convergent then, for large enough k, for any current strategy profile $s^t \in S$, level k logit choice by player i will choose the Nash equilibrium strategy $B_i^\infty(s^t)$ with probability approaching one as $\eta \to 0$.

3.4 Dynamic strategy updating

We define the *level k logit dynamics* on the state space of strategy profiles. The game is played repeatedly in discrete time and strategies are updated according to the level k logit choice rule. Let the strategy profile played at period t be s^t. At time $t+1$, any given player i is, independently of the other players, *active* with probability $q_i \in (0,1)$. If i is not active at $t+1$, then his strategy at $t+1$ remains the same as his strategy at t. That is, $s_i^{t+1} = s_i^t$. If player i is active at period $t+1$ and is of level k, then he updates his strategy according to the level k logit choice rule.

Remark 3. If $k=1$ for every player, then this process is effectively the standard logit dynamic of Blume (1993).[5] Players for whom $k=2$ correspond to the 'clever' players of Sáez-Martı and Weibull (1999); Matros (2003).

Remark 4. Let G be Nash convergent and every player have a level k high enough that $B^{k-1}(\cdot) = B^\infty(\cdot)$. Under the unperturbed ($\eta = 0$) dynamic, if the strategy profile at time t is s^t, then with probability at least $\prod_{i\in N} q_i$, the strategy profile at time $t+1$ will be the Nash equilibrium $B^\infty(s^t)$. High rationality players that achieve such coordination correspond to the 'sophisticated' players of Myatt and Wallace (2003).

Remarks 3 and 4 illustrate that the level k logit dynamics bridge the gap between perturbed best response dynamics in the style of Kandori et al. (1993); Young (1993a) and instantaneously jumping to a Nash equilibrium. Indeed, an important implication of the current paper is that certain results are robust across this entire class of models.

[3]Note that players' conjectures do not consider the possibility that other players' choices are perturbed. This is in contrast to, for example, the 'noisy introspection' model of Goeree and Holt (2004). For more general conjectures about the behavior of others, see Section 6.

[4]In fact, it can be checked that, given current strategy profile s^t, logit choice probabilities under a sequence of decreasing values of η correspond to a sequence that identifies $B^\infty(s^t)$ as a proper equilibrium under the definition of Myerson (1978).

[5]See Alós-Ferrer and Netzer (2010) for an extended discussion of this process and Sandholm (2010); Newton (2018) for discussion of related processes.

3.5 Stochastic stability

Under the level k logit dynamics with $\eta > 0$, any state can be reached from any other state. Therefore the process is irreducible and has a unique stationary distribution, which we denote π_η. The stationary distribution gives the proportion of time that the process will spend at any strategy profile in the long run. We are interested in dynamics that are close to best response dynamics, that is when η is small.

It can be shown by standard arguments (Foster and Young, 1990) that π_η converges as $\eta \to 0$. Denote this limiting stationary distribution by π_0. If $\pi_0(s) > 0$, we say that s is *stochastically stable*. If $\pi_0(s) = 1$, we say that s is *uniquely stochastically stable*. For small values of η, the process will spend almost all of its time at stochastically stable strategy profiles. Thus the identity of stochastically stable states tells us which strategy profiles we can expect to be played most of the time in the long run.

4. Main result

Our main result is that results on the stability of risk dominant strategy profiles under standard ($k = 1$) perturbed best response dynamics are robust to level k updating. We shall withhold discussion of why the result is novel, interesting and non-obvious until after presenting it. First, we shall define the concept of risk dominance that we use, *generalized risk dominance* (Peski, 2010).

Consider any given strategy profile and label it s^A. Without loss of generality, we label the strategies of every player at s^A as A, so that $s_i^A = A$ for all $i \in N$. If a pair of strategy profiles s, s' are such that every player plays A at at least one of s and s', then we say that s and s' are *A-associated* (see Figure 3). Generalized risk dominance of s^A holds when, for any two A-associated strategy profiles, the incentives to play A at one of these profiles outweighs any incentive not to play A at the other profile.[6]

Definition 1 (Generalized risk dominance)**.**
Profile s^A is *generalized risk dominant* (GR-dominant) if, for all A-associated strategy profiles s', s'', for all $i \in N$,

$$(5) \qquad U_i(A, s'_{-i}) + U_i(A, s''_{-i}) \geq \max_{s_i \neq A} U_i(s_i, s'_{-i}) + \max_{s_i \neq A} U_i(s_i, s''_{-i}).$$

Substituting $>$ for $\geq$ in (5) gives the definition of *strict generalized risk dominance*.

[6]Peski (2010) defines ordinal and cardinal forms of GR-dominance. The definition we use corresponds to the cardinal form.

	s^A	s	s'	s^B
Alice	A	A	B	B
Bob	A	B	A	B

Figure 3: A**-association.** Consider the illustrated example with two players and two strategies, labeled A and B, for each player. It can be seen that s^A is A-associated with every other strategy profile. Furthermore, s is A-associated with s', as every player plays A at at least one of these two profiles. However, s is not A-associated with s^B, as Bob plays A at neither of these profiles. Finally, s' and s^B are not A-associated, as Alice plays A at neither of these profiles.

Remark 5. With two players and two strategies, (strict) GR-dominance is equivalent to (strict) risk dominance of Harsanyi and Selten (1988). We consider this further in Section 5. For two players and more than two strategies, (strict) GR-dominance implies (strict) $\frac{1}{2}$-dominance of Morris et al. (1995). Conversely, when there are many players but payoffs are a linear sum of payoffs from two player interactions, a strong form of $\frac{1}{2}$-dominance implies GR-dominance (Peski, 2010). In general, however, these concepts are independent (see also Iijima, 2015).

We are now ready to state our main Theorem. Risk dominance is robustly selected for under the entire class of level k logit dynamics.

Theorem 1. *Under the level k logit dynamics*

- *If s^A is GR-dominant, then s^A is stochastically stable.*
- *If s^A is strictly GR-dominant, then s^A is uniquely stochastically stable.*

Remark 6. It is known that stochastic stability of GR-dominant profiles holds under a class of payoff dependent behavioral rules (see Section 6) that includes logit (Peski, 2010). This is a generalization of earlier results concerning the stochastic stability of risk dominant profiles in 2 by 2 games (Kandori et al., 1993; Young, 1993a; Blume, 1993, 2003). It is further known that these results are robust to heterogeneity in behavioral rules (Newton, 2020). However, unlike the above work, the current work considers non-myopic conjectures. Specifically, the above papers consider smoothed best responses to the current strategy profile s^t, whereas we consider smoothed best responses to a variety of conjectures based on s^t. Furthermore, we allow heterogeneity in these conjectures across players.

Remark 7. The seminal papers of Kandori et al. (1993); Young (1993a); Blume (1993) can be considered to have made two principal contributions. (I) The unperturbed ($\eta = 0$) dynamic eventually converges to a Nash equilibrium with probability one (under conditions of what Young calls *weak acyclicity*), and (II) in two strategy coordination games, risk dominant Nash equilibria are stochastically stable. In the current model, (I) does not hold. Persistent miscoordination may arise due to $k > 1$ and can prevent convergence of the unperturbed dynamic to a Nash equilibrium (see Section 5 for an example). Nevertheless, result (II) continues to hold. That is, the long run stability of risk dominance does not rely on short run convergence to Nash equilibrium.

Remark 8. Consider Nash convergent G, small η and q_i close to 1. Under these conditions, from any non-Nash equilibrium profile, the process will reach a Nash equilibrium in a single period with a probability close to 1. This instant convergence, often assumed in one-shot games, does not change the stochastic stability of risk dominance. Taken together with Remark 7, this implies that stochastic stability of risk dominance is unaffected by either of the contrasting cases of instant convergence or non-convergence of the unperturbed dynamic to Nash equilibrium.

Remark 9. We show in Section 6 that our results our robust to the levels of players being generated randomly in each period in a way that allows for correlation, both positive and negative, across players. Consider an alternative approach of adding a state variable that tracks players' levels, with levels increasing over time. Considering this state variable as part of the state space, the process is no longer irreducible. However, if G is Nash convergent, then the process governing evolution of the strategy profile converges as best responses converge to $B^{\infty}(\cdot)$. This implies that the behavioral implications of Theorem 1 continue to hold.

Sketch of Proof. The proof of Theorem 1 is given in Appendix A. A summary of the proof is as follows. First, we disaggregate the process and consider processes in which only a single player of some given level k updates his strategy, with the strategies of other players remaining fixed. Note that such processes are not irreducible. However, this does not matter, as we can still show that they satisfy a particular property. Specifically, we show in Lemma 2 that if s^A is GR-dominant, then these processes satisfy a form of asymmetry towards strategy A. This form of asymmetry was considered for processes by Peski (2010) and applied to individual behavioral rules by Newton (2020). Second, we re-aggregate the process to once again consider the process in which every player updates independently as described in our model. We show in Lemma 3 that asymmetry of the disaggregated processes implies asymmetry of the aggregate process. This is done by applying Theorem 3 of Newton (2020, 2019). Finally, we apply Theorem 1 of Peski (2010), which states that asymmetry of the aggregate process towards strategy A implies stochastic stability of s^A.

We shall end this section with a brief consideration of the possible impact of players who play completely randomly. Specifically, let a *level* 0 *player* be a player who chooses

	A	B
A	a_{AA}	a_{AB}
B	a_{BA}	a_{BB}

Figure 4: **Unweighted payoffs from each interaction.** $a_{AA} > a_{BA}$ and $a_{BB} > a_{AB}$. For each combination of A and B, entries give payoffs for the row player.

a strategy uniformly at random from his strategy set. That is, if player i is a level 0 player, then when he updates his strategy he will choose each strategy in S_i with probability $1/|S_i|$. In the presence of such players, the first part of Theorem 1 continues to hold. However, the second part of Theorem 1 does not. Even under strict GR-dominance of s^A, there are always multiple stochastically stable states when any level 0 player has more than one strategy.

Theorem 2. *Let every player either follow the level k logit dynamics for $k \geq 1$ or be a level 0 player. If s^A is GR-dominant, then s^A is stochastically stable. If there exists a level 0 player with more than one strategy, then there exist multiple stochastically stable strategy profiles.*

The extent of the multiplicity caused by level 0 players is sensitive to model details. In particular, it can depend on the levels of the non-level 0 players, as we will see in Example 2 of Section 5.

5. Application: technology adoption on networks

Consider a situation in which each player may adopt one of two technologies. Specifically, let $S_i = \{A, B\}$ for all $i \in N$. Each player is influenced by other players and may be influenced by some players more than others. Let $w_{ij} \in \mathbb{R}_{\geq 0}$ be the influence of player j on player i. Assume that every player i is influenced by at least one other player, so that $\sum_{j \in N \smallsetminus \{i\}} w_{ij} > 0$, as is the case in the game illustrated in Figure 5. Each player wishes to adopt a similar strategy to those who influence him. Specifically, payoffs from each pairwise interaction are given by the game illustrated in Figure 4. The payoff to player i at strategy profile s is then the sum of the payoffs from his pairwise interactions weighted by their influence. That is,

$$(6) \qquad U_i(s) = U_i(s_i, s_{-i}) = \sum_{j \in N \smallsetminus \{i\}} w_{ij} a_{s_i s_j},$$

where $a_{s_i s_j} \in \mathbb{R}$ is the payoff to player i from his interaction with player j.

A classic result (Blume, 1993) is the stochastic stability of strategy profiles at which every player plays the same risk dominant strategy (Harsanyi and Selten, 1988). In the two player game of Figure 4, a strategy is risk dominant if it maximizes payoff given that the opposing player randomizes evenly across his two strategies.

Definition 2.
Strategy A is *risk dominant* if

$$a_{AA} + a_{AB} \geq a_{BA} + a_{BB}, \tag{7}$$

and is *strictly risk dominant* if the inequality holds strictly.

It is possible to show that (strict) risk dominance of A implies (strict) GR-dominance of s^A. We can then apply Theorem 1 to show that (strict) risk dominance of A implies (unique) stochastic stability of s^A. The reverse implication then follows from the fact that at least one of the two strategies must be risk dominant.

Proposition 1. *For technology adoption on a network under the level k logit dynamics,*

- s^A *is stochastically stable if and only if A is risk dominant.*
- s^A *is uniquely stochastically stable if and only if A is strictly risk dominant.*

The special case of Proposition 1 in which all players are level 1 is known from Blume (1993, 1996). Proposition 1 shows that stochastic stability of risk dominance is robust to varying levels of rationality. The result is not obvious. Level k thinking can lead to short run behavior that is completely different to that implied by level 1 thinking. However, in all cases, long run behavior tends to risk dominance (see Remark 7 earlier in the paper). We illustrate these points with the following example.

Example 1. Let $N = \{i, j\}$ and $w_{ij} = w_{ji} = 1$. If player i is of level 2, then he will never change his strategy as a result of playing a best response to his conjecture. To see this, let the current strategy profile be $s^t = (s_i^t, s_j^t)$. Given this current strategy profile, player i will conjecture that player j will play $B_j^1(s^t) = s_i^t$. That is, player i expects player j at time $t+1$ to coordinate with the strategy of player i at time t. A best response for player i to this conjecture is to remain playing the same strategy at time $t+1$ as he plays at time t. That is, he does not change his strategy. Consequently, if both player i and player j are level 2, then neither of them will ever change his strategy as a result of a best response. It follows that all strategy profiles are absorbing states of the process with $\eta = 0$. This is in stark contrast to the standard case in which every player has level 1, where the process with $\eta = 0$ converges with probability one in finite time to a Nash equilibrium of the game. Nevertheless, Proposition 1 tells us that level k does not affect stochastic stability predictions for the perturbed process. The stability of risk dominance is robust to rationality.

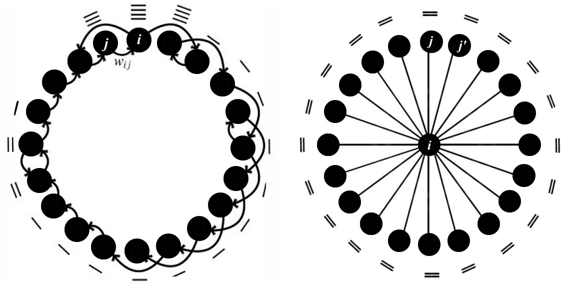

Figure 5: **Influence networks.** Nodes are players. Dashes adjacent to nodes indicate players' levels of rationality, here between zero (no line) and four. **Left:** Example cognitive hierarchy and influence structure. Directed edges indicate non-zero influences and their direction. That is, an arrow from i to j indicates that the influence of player j on player i (i.e. w_{ij}) is strictly positive. **Right:** Star network from Example 2 ('even' case). Links indicate mutual influence between two players. Per Example 2, $w_{ik} = w_{ki} = 1$ for all $k \neq i$.

Theorem 2 showed that the presence of level 0 players who act randomly and unpredictably can lead to multiplicity of stochastically stable strategy profiles even when s^A is strictly GR-dominant. The following example shows how the extent of this multiplicity can be sensitive to the levels of the non-level 0 players.

Example 2. Let there be $|N| \geq 3$ players interacting on a star network (as illustrated on the right of Figure 5). That is, there is is some player i such that $w_{ij} = w_{ji} = 1$ for all $j \neq i$, and $w_{jk} = 0$ for all $j \neq i$, $k \neq i$. Let player i be a level 0 player and let all $j \neq i$ be level k^*. Let A be strictly risk dominant, hence s^A is strictly GR-dominant by the proof of Proposition 1.

The first part of Theorem 2 implies that s^A is stochastically stable. The second part of Theorem 2 implies that if there is at least one level 0 player, then there must also exist stochastically stable profiles other than s^A. Somewhat trivially, if s^A is stochastically stable, then the profile at which player i plays B and all other players play A must be stochastically stable. However, the randomness in the behavior of player i may also be enough to change the behavior of the other players.

Proposition 2. *Consider Example 2. If k^* is odd, then all strategy profiles are stochastically stable. If k^* is even, then there are two stochastically stable strategy profiles, $(s_i, s_{-i}) = s^A$ and $(s_i, s_{-i}) = (B, s^A_{-i})$.*

To see the intuition behind Proposition 2, first consider $k^* = 1$. Starting from s^A, if player i switches to B, then any other player j may subsequently switch to B as a best response. This introduces enough disorder into the system that every strategy profile can be reached by the unperturbed dynamic ($\eta = 0$) and, as a consequence, every strategy profile is stochastically stable. All odd values of k^* follow similar reasoning.

Now consider $k^* = 2$. From any initial strategy profile s, a player $j \neq i$ will conjecture that player i will play $B_i^1(s)$ in the next period. Note that $B_i^1(s)$ does not depend on s_i, therefore the conjecture that j makes about the behavior of player i does not depend on player i's current strategy. Therefore, the behavior of player j does not depend on player i's current strategy. Consequently, the long run behavior of j is independent of the behavioral rule used by i. Consider the hypothetical case in which player i has level $k \geq 1$. We know by Theorem 1 that s^A would then be uniquely stochastically stable. In particular, player j would play A most of the time in the long run. As the behavior of player j is independent of the behavior of player i, this must also be true when player i has level 0. Therefore, at any stochastically stable profile it must be that all $j \neq i$ play A. That is, s^A and (B, s^A_{-i}) are the only stochastically stable profiles. All even values of k^* follow similar reasoning.

6. Generalization

In this section we generalize the model in several dimensions. As before, we consider players who do not best respond to the current strategy profile, but rather form conjectures about play at $t+1$ to which they best respond. However, now we do not restrict attention to level k conjectures, but rather consider a more general class. In general, a *conjecture* for player i is a function $f_i : S \to S$. A profile of conjectures is given by $f = \{f_i\}_{i \in N}$.

We will show that an important class of conjectures are those that preserve A-association. Given A-associated profiles s, s', this requires that the respective conjectures formed when presented with these strategy profiles are themselves A-associated.

Definition 3 (A-association preserving)**.** Profile of conjectures f *preserves A-association* if, for all $i \in N$, s, s' A-associated, we have that $f_i(s), f_i(s')$ are A-associated.

The conjectures considered so far in the paper correspond to $f_i(s) = s$ and $f_i(s) = B^k(s)$ for $k \geq 1$. The conjecture $f_i(s) = s$ always satisfies Definition 3 (see below). For $f_i(s) = B^k(s)$, if s^A is GR-dominant and best responses are unique then Definition 3 is satisfied. Given that this was our only use of the assumption of unique best responses, for the remainder of the paper we drop this assumption.

We present some examples of conjectures that satisfy Definition 3 irrespective of whether GR-dominance holds.

Example 3 (Myopia)**.** Consider the conjecture that all players remain playing their current strategy. That is, $f_i(s) = s$. It follows trivially that if s, s' are A-associated, then $f_i(s), f_i(s')$ are A-associated.

Example 4 (Majoritarianism)**.** Let $|N|$ be odd and consider the conjecture that all players play the most popular current strategy, with some tie breaking rule employed. If s, s' are A-associated, it must be that a majority of players at s or s' play A. Consequently, $f_i(s) = s^A$ or $f_i(s') = s^A$, therefore $f_i(s), f_i(s')$ are A-associated.

Example 5 (Imitate a friend)**.** Consider the conjecture that each player imitates some other player. That is, for all $j \in N$, we have $(f_i(s))_j = s_k$ for some $k \in N$. If s, s' are A-associated, then for all k, $s_k = A$ or $s'_k = A$. Consequently, for all j, $(f_i(s))_j = A$ or $(f_i(s'))_j = A$. Therefore, $f_i(s), f_i(s')$ are A-associated.

Another dimension that we generalize is the behavioral rule that acts on the conjectured strategy profile. So far, we have considered the logit choice rule. Now, we consider a large class of *regular* behavioral rules (Young, 1993a) in which behavior depends on payoff differences.[7]

Adjust the behavioral rule given in Section 3 as follows. If player i is active at period $t+1$, then he forms a conjecture $f_i(s^t)$ about behavior at time $t+1$. Given his conjecture about the behavior of other players, he then updates his strategy according to a perturbed best response rule parameterized by η. Let $\Upsilon_i(s'_i, s'')$ denote the expected payoff loss, relative to a best response, incurred by player i when he plays s'_i against s''_{-i}. That is,

$$\Upsilon_i(s'_i, s'') = \max_{s_i \in S_i} U_i\left(s_i, s''_{-i}\right) - U_i\left(s'_i, s''_{-i}\right) \tag{8}$$

The probability that $s_i^{t+1} = s'_i$ at time $t+1$ is then given by

$$Pr\left(s_i^{t+1} = s'_i | s^t\right) = \Big(a + o(1)\Big) e^{-\frac{1}{\eta} g_i\left(\Upsilon_i\left(s'_i, f_i(s^t)\right)\right)}, \tag{9}$$

where a is a strictly positive constant that can depend on s'_i, s^t, f_i, but not on η; $o(1)$ is a term that approaches zero as $\eta \to 0$; g_i is a nonnegative, weakly increasing function.

Note that larger values of $g^i(\cdot)$ in (9) imply smaller probabilities. Together with (8), this implies that the probability of choosing s'_i decreases in the difference between the payoff from best responding to the conjectured profile $f_i(s^t)$ and the payoff from playing s'_i against $f_i(s^t)$. The best response plus uniform deviations rule corresponds to $g_i(x) = \text{sgn}(x)$ and the logit choice rule corresponds to $g_i(x) = x$ for appropriate choice of a and $o(1)$.

We are now ready to present our generalization of Theorem 1. If conjectures preserve A association, then under the class of perturbed best response rules described by (9), GR-dominance implies stochastic stability. Thus, the results of the current paper extend to a wide

[7] These behavioral rules roughly correspond to *skew-symmetric* rules (Blume, 2003), *payoff-based* rules (Peski, 2010) and *self regarding payoff-difference based* rules (Newton, 2020).

range of conjectures and behavioral rules.

Theorem 3. *Let f preserve A-association. Under the dynamics (9)*

- *If s^A is GR-dominant, then s^A is stochastically stable.*
- *If s^A is strictly GR-dominant and, for all $i \in N$, we have g_i strictly increasing and $f_i(s^A) = s^A$, then s^A is uniquely stochastically stable.*

We present one final generalization. It may be that player's conjectures vary from period to period. It may even be the case that player's conjectures are correlated with each other. For example, it may be that a player is more likely to exhibit low rationality behavior when other players are exhibiting low rationality behavior. It turns out that randomness and correlation in conjectures does not affect our results.

Let F be a set of profiles of conjectures and let φ be an exogenously given distribution over F. Adjust the model so that rather than there being a single fixed profile of conjectures, every period a profile of conjectures f is chosen according to φ.

Theorem 4. *Let all $f \in F$ preserve A-association. Under the dynamics (9)*

- *If s^A is GR-dominant, then s^A is stochastically stable.*
- *If s^A is strictly GR-dominant and, for all $i \in N$, we have g_i strictly increasing and $f_i(s^A) = s^A$ for all $f \in F$, then s^A is uniquely stochastically stable.*

7. Application: Second price auctions

Consider a second price auction in which players bid for ownership of some good. Let $V_i \subset \mathbb{R}$, $|V_i| \geq 2$, be a finite set of possible valuations that player i may have for the good. Let v_i denote a representative element of V_i. Similarly write $V = \times_{i \in N} V_i$ and denote a representative element of V by $v = (v_i)_{i \in N}$. That is, v denotes a possible profile of valuations. The finite set of possible bids that player i can make is denoted $B_i \supseteq V_i$. The set of strategies S_i of player i is then the finite set of functions that map valuations to bids. That is, $s_i \in S_i$ if and only if s_i is a function $s_i : V_i \to B_i$. Given strategy profile s and valuations v, the payoff of player i is given by

$$u_i(s,v) := \begin{cases} \frac{v_i - \max_{j \neq i} s_j(v_j)}{|\arg\max_{j \in N} s_j(v_j)|} & \text{if } i \in \arg\max_{j \in N} s_j(v_j), \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

That is, the good is given to one of the players amongst those with the highest bids. If several players make the same highest bid then one of them is chosen uniformly at random to receive the good. The player who receives the good gets a payoff equal to the difference between his valuation and the second highest bid. All other players pay nothing, do not obtain the good, and therefore receive a payoff of zero. Letting $\mathscr{V}$ be an exogenously given distribution over valuations V, we then have ex-ante expected payoffs given by

$$U_i(s) = \mathbb{E}_{\mathscr{V}(\cdot)} u_i(s, \cdot). \tag{11}$$

Before we proceed to analyze this problem, we require a further general result. Nash equilibria in weakly dominant strategies are GR-dominant and Nash equilibria in strictly dominant strategies are strictly GR-dominant. Strategy A is *weakly dominant* when for all $s'_i \neq A$, we have $U_i(A, s_{-i}) \geq U_i(s'_i, s_{-i})$ for all $s_{-i} \in S_{-i}$.[8] Strategy A is *strictly dominant* when for all $s'_i \neq A$, we have $U_i(A, s_{-i}) > U_i(s'_i, s_{-i})$ for all $s_{-i} \in S_{-i}$.

Theorem 5. *If s^A is composed of weakly dominant strategies, then s^A is GR-dominant. If s^A is composed of strictly dominant strategies, then s^A is strictly GR-dominant.*

For each player i, let s_i^A be such that $s_i^A(v_i) = v_i$ for all $v_i \in V_i$. That is, A is the strategy according to which the bid of player i always equals his valuation. We refer to s^A as *bidding one's own valuation*. Note that A is a weakly dominant strategy. Combining Theorem 5 with Theorem 4, we therefore conclude that bidding one's valuation in the second price auction is stochastically stable under our general dynamics.

Proposition 3. *When the second price auction is played under the dynamics (9) and all $f \in F$ preserve A-association, then bidding one's own valuation is stochastically stable.*

Note that the best response correspondence in this game is multi-valued, so there is no guarantee that a conjecture based on best responses will preserve A-association, even when s^A is GR-dominant. However, if we restrict attention to best responses such that if s_j is a best response to s, then player j remains playing s_j, then we do indeed preserve A-association. This is easily verified using the definition of A-association and the fact that weak dominance of A implies that $A \in B_j^1(s)$.

Under some further assumptions we can obtain unique stochastic stability of bidding one's own valuation. Call a strategy $s_i \in S_i$ *sensible* if $s_i(v_i) \leq v_i$ for all $v_i \in V_i$. That is, a sensible strategy is a strategy according to which a player never bids strictly more than his valuation. Restrict attention to sensible strategies, to $B_i = V_i$ and to valuations v_i which are

[8] Note that this is weaker than the usual definition of weak dominance which requires in addition that $U_i(A, s_{-i}) > U_i(s'_i, s_{-i})$ for some $s_{-i} \in S_{-i}$.

independent, identically distributed (*iid*) random variables with full support on V_i. Consider a strategy $s_i^* \neq A$. As s_i^* is sensible and $s_i^* \neq A$, there must be some $v_i^* \in V_i$ such that $s_i^*(v_i^*) < v_i^*$. Let us compare s_i^* to A.

First, note that if all opposing players play sensible strategies and at least one (say player j) of the opposing players plays A, then i will sometimes obtain a strictly lower realized payoff when playing s_i^* than he would have obtained by playing A. To see this, consider the case when i has valuation v_i^* and all other players have valuation $s_i^*(v_i^*)$. Player j follows strategy A and so will bid $s_i^*(v_i^*)$ at these valuations. All players $k \neq i, j$ play sensible strategies and so will bid no more than $s_i^*(v_i^*)$. Consequently, when player i follows strategy s_i^*, his bid will equal player j's bid so that sometimes player i will fail to win the good and will obtain a realized payoff of zero. Had player i instead followed strategy A, he would have bid $v_i^* > s_i^*(v_i^*)$, won the good for sure and obtained a realized payoff of $v_i^* - s_i^*(v_i^*) > 0$.

Second, note that if strategy profiles s' and s'' are A-associated, then either $s'_j = A$ or $s''_j = A$. Combining these observations and assuming without loss of generality that $s'_j = A$, we have

$$(12) \qquad U_i(A, s'_{-i}) > U_i(s_i^*, s'_{-i}).$$

Next, note that weak dominance of A implies that

$$(13) \qquad U_i(A, s''_{-i}) \geq U_i(s_i^*, s''_{-i}).$$

Summing (12) and (13), we obtain the condition for strict GR-dominance in Definition 1. Again applying Theorem 4, we see that bidding one's own valuation is uniquely stochastically stable under these conditions.

Proposition 4. *Let $B_i = V_i$ for all i, valuations v_i be iid with full support on V_i and strategies be sensible. When the second price auction is played under the dynamics (9) with g_i strictly increasing for all $i \in N$, and all $f \in F$ preserve A-association and satisfy $f_i(s^A) = s^A$ for all $i \in N$, then bidding one's own valuation is uniquely stochastically stable.*

In summary, bidding one's own valuation in the second price auction is robust to a large range of behavioral rules. Whether players are low rationality or high rationality, whether they are homogeneous or diverse in their levels of rationality, the stability of bidding one's own valuation is preserved.

This concludes the main body of the paper.

Appendix

A. Proofs

A.1 Additional definitions and useful results

For parameter value η, strategy profiles $s, s' \in S$, let $P^{\eta}(s, s')$ denote the probability that $s^{t+1} = s'$ conditional on $s^t = s$.

Define a new Markov process on S, denoted P_i^{η}, by adjusting the original process by setting $q_j = 0$ for all $j \neq i$. Let $P_i^{\eta}(s, s')$ denote the probability that $s^{t+1} = s'$ conditional on $s^t = s$. Observe that, for all $\eta > 0$, $s, s' \in S$, we have

$$(14) \qquad P^{\eta}\left(s, s'\right) = \prod_{i \in N} P_i^{\eta}\left(s, (s_i', s_{-i})\right).$$

Define cost functions

$$(15) \qquad c_i(s, s') := \begin{cases} \lim_{\eta \to 0} -\eta \log P_i^{\eta}(s, s') & \text{if } P_i^{\hat{\eta}}(s, s') > 0 \text{ for some } \hat{\eta} > 0, \\ \infty & \text{otherwise .} \end{cases}$$

and let $c(s, s')$ be the equivalent quantity after replacing P_i^{η} by P^{η}.

Simple algebra shows that, for the updating rule in our model, we have

$$(16) \qquad c_i(s, s') := \begin{cases} 0 & \text{if } s' = s, \\ \max_{x_i \in S_i} U_i\left(x_i, B_{-i}^{k-1}(s)\right) - U_i\left(s_i', B_{-i}^{k-1}(s)\right) & \text{if } s_i' \neq s_i,\, s_{-i}' = s_{-i}, \\ \infty & \text{otherwise .} \end{cases}$$

We adopt the convention that $\infty > \infty$.

We say that s'' A-dominates s' if $s_i'' = A$ for all i such that $s_i' = A$.

Definition 4. $c(\cdot, \cdot)$ is *asymmetric* (towards A) if, for any $s, s', \tilde{s}$ such that $s, \tilde{s}$ are A-associated, there exists $\tilde{s}'$ such that

- $\tilde{s}'$ A-dominates $\tilde{s}$,
- $s', \tilde{s}'$ are A-associated, and
- $c(s, s') \geq c(\tilde{s}, \tilde{s}')$.

Definition 5. $c(\cdot, \cdot)$ is *strictly asymmetric* (towards A) if

(a) for any $s \neq s^A$, $c(s^A, s) > 0$, and

(b) for any $s, s', \tilde{s}$ such that $s, \tilde{s}$ are A-associated, there exists $\tilde{s}'$ such that

- $\tilde{s}'$ A-dominates $\tilde{s}$,
- $s', \tilde{s}'$ are A-associated, and
- either $c(\tilde{s}, \tilde{s}') = 0$ or $c(s, s') > c(\tilde{s}, \tilde{s}')$.

Note that Definitions 4 and 5 can be applied to both $c(\cdot,\cdot)$ and to $c_i(\cdot,\cdot)$.

Theorem N (Newton, 2020, 2019)**.**
Let P^η, $\{P_i^\eta\}_{i\in N}$ satisfy (14). ***(i)*** *If $c_i(\cdot,\cdot)$ is asymmetric for all $i \in N$, then $c(\cdot,\cdot)$ is asymmetric (Newton, 2020, Theorem 3).* ***(ii)*** *If $c_i(\cdot,\cdot)$ is strictly asymmetric for all $i \in N$, then $c(\cdot,\cdot)$ is strictly asymmetric (Newton, 2019, Theorem 3).*

Theorem P (Peski, 2010, Theorem 1)**.**
(i) *If $c(\cdot,\cdot)$ is asymmetric, then s^A is stochastically stable.* ***(ii)*** *If $c(\cdot,\cdot)$ is strictly asymmetric, then s^A is uniquely stochastically stable.*

A.2 Proofs for Section 4

Lemma 1. *Let $s, \tilde{s}$ be A-associated. If s^A is GR-dominant, then $B^k(s), B^k(\tilde{s})$ are A-associated for all $k \geq 1$.*

Proof. If $s, \tilde{s}$ are A-associated, then by GR-dominance of s^A, in particular expression (5), we have

$$U_i(A, \tilde{s}_{-i}) - \max_{x_i \neq A} U_i(x_i, \tilde{s}_{-i}) \geq \max_{x_i \neq A} U_i(x_i, s_{-i}) - U_i(A, s_{-i}). \tag{17}$$

If $B_i^1(s) \neq A$, we have

$$\max_{s_i \neq A} U_i(s_i, s_{-i}) - U_i(A, s_{-i}) > 0, \tag{18}$$

so combining (17) and (18) we obtain

$$U_i(A, \tilde{s}_{-i}) - \max_{x_i \neq A} U_i(x_i, \tilde{s}_{-i}) > 0. \tag{19}$$

Therefore, $B_i^1(\tilde{s}) = A$. This holds for all i such that $B_i^1(s) \neq A$, therefore $B_i^1(s), B_i^1(\tilde{s})$ are A-associated. Iterating the above argument, we obtain that $B_i^k(s), B_i^k(\tilde{s})$ are A-associated for $k = 2, 3, \ldots$. □

Lemma 2. ***(i)*** *If s^A is GR-dominant, then, for all $i \in N$, c_i is asymmetric towards A.* ***(ii)*** *If s^A is strictly GR-dominant, then, for all $i \in N$, c_i is strictly asymmetric towards A.*

Proof. Note that as s^A is A-associated with itself, GR-dominance of s^A and uniqueness of best responses implies that

$$(20) \qquad B^k(s^A) = s^A \quad \text{for all } k \geq 0.$$

If $s \neq s^A$, then either $s_{-i} \neq s^A_{-i}$, in which case $c_i(s^A, s) = \infty$, or $s_{-i} = s^A_{-i}$, $s_i \neq A$, in which case

$$(21) \qquad \begin{aligned} c_i(s^A, s) &\underbrace{=}_{\text{by (16)}} \max_{x_i \in S_i} U_i(x_i, B^{k-1}_{-i}(s^A)) - U_i(s_i, B^{k-1}_{-i}(s^A)) \\ &\underbrace{=}_{\text{by (20)}} \max_{x_i \in S_i} U_i(x_i, s^A) - U_i(s_i, s^A) \\ &= U_i(A, s^A) - U_i(s_i, s^A) \\ &> 0. \end{aligned}$$

Therefore, the condition in Definition 5a is satisfied.

Now consider $s, s', \tilde{s} \in S$ such that $s, \tilde{s}$ are A-associated. If s^A is GR-dominant, it follows from Lemma 1 that

$$(22) \qquad B^k(s), B^k(\tilde{s}) \quad \text{are } A\text{-associated for all } k \geq 1.$$

<u>**Case 1**</u>**:** $s = s'$ or $s'_i = A$ or $c_i(s, s') = \infty$ or $\tilde{s}_i = A$.

If $c_i(s, s') = \infty$, let $\tilde{s}' = s^A$. The conditions in Definitions 4 and 5b are satisfied.

If $c_i(s, s')$ is finite, let $\tilde{s}' = \tilde{s}$. (16) implies $c_i(\tilde{s}, \tilde{s}') = 0$, therefore the conditions in Definitions 4 and 5b are satisfied.

<u>**Case 2**</u>**:** $s \neq s'$ and $s'_i \neq A$ and $c_i(s, s')$ is finite and $\tilde{s}_i \neq A$.

(16) together with finiteness of $c_i(s, s')$ implies $s_{-i} = s'_{-i}$. $s \neq s'$ and $s'_i \neq A$ then imply that $s_i = A$. Let $\tilde{s}'$ be such that $\tilde{s}'_{-i} = \tilde{s}_{-i}$, $\tilde{s}'_i = A$.

If $c_i(\tilde{s}, \tilde{s}') = 0$, then the conditions in Definitions 4 and 5b are satisfied.

If $c_i(\tilde{s}, \tilde{s}') > 0$, then

$$(23) \qquad \begin{aligned} c_i(s, s') &\underbrace{=}_{\text{by (16)}} \max_{x_i \in S_i} U_i(x_i, B^{k-1}_{-i}(s)) - U_i(s'_i, B^{k-1}_{-i}(s)) \\ &\geq U_i(A, B^{k-1}_{-i}(s)) - \max_{x_i \neq A} U_i(x_i, B^{k-1}_{-i}(s)) \\ &\underbrace{\geq}_{\substack{\text{by (22)} \\ \text{and GR-dominance}}} \max_{x_i \neq A} U_i(x_i, B^{k-1}_{-i}(\tilde{s})) - U_i(A, B^{k-1}_{-i}(\tilde{s})) \\ &\underbrace{=}_{\substack{\text{by (16)} \\ \text{and } c(\tilde{s}, \tilde{s}') > 0}} c_i(\tilde{s}, \tilde{s}'). \end{aligned}$$

That is, the condition in Definition 4 as satisfied. Further note that if s^A is strictly GR-dominant, then the weak inequality in (23) due to GR-dominance becomes a strict inequality, so that the condition in Definition 5b is satisfied. □

Lemma 3. ***(i)*** *If s^A is GR-dominant, then c is asymmetric towards A.* ***(ii)*** *If s^A is strictly GR-dominant, then c is strictly asymmetric towards A.*

Proof. **(i)** GR-dominance of s^A and Lemma 2(i) together imply that, for all $i \in N$, c_i is asymmetric towards A. Theorem N(i) then implies that c is asymmetric towards A. **(ii)** Strict GR-dominance of s^A and Lemma 2(ii) together imply that, for all $i \in N$, c_i is strictly asymmetric towards A. Theorem N(ii) then implies that c is strictly asymmetric towards A. □

Proof of Theorem 1.
Assume s^A is GR-dominant. Lemma 3(i) and Theorem P(i) together imply stochastic stability of s^A.

Assume s^A is strictly GR-dominant. Lemma 3(ii) and Theorem P(ii) together imply unique stochastic stability of s^A. □

Proof of Theorem 2.
Let player i be level 0. Recall that when i updates his strategy, each strategy in S_i is chosen with probability $1/|S_i|$. This probability is independent of η and s_{-i}, so (15) gives us $c_i(s_i,(s'_i,s_{-i})) = 0$ for all s, s'_i. This c_i satisfies Definition 4, so c_i is asymmetric.

Therefore, even if some players are level 0, Lemma 3(i) and the first part of Theorem 1 continue to hold. Therefore, if s^A is GR-dominant, then s^A is stochastically stable.

To prove the second part of the Theorem, let i have at least two strategies. Let s^* be stochastically stable. Consider $s'_i \in S_i$ such that $s'_i \neq s^*_i$. Then $c_i(s^*,(s'_i,s^*_{-i})) = 0$, therefore $c(s^*,(s'_i,s^*_{-i})) = 0$. It follows from the tree characterization of stochastically stable states (see, e.g. Young, 1993a) that (s'_i,s^*_{-i}) must also be stochastically stable. □

A.3 Proofs for Section 5

Proof of Proposition 1.

Step 1

First we show that if A is (strictly) risk dominant, then s^A is (strictly) GR-dominant. Note that condition (5) for GR-dominance reduces to

$$U_i(A,s'_{-i}) + U_i(A,s''_{-i}) \geq U_i(B,s'_{-i}) + U_i(B,s''_{-i}) \tag{24}$$

for all $i \in N$, s', s'' A-associated.

Now, if s', s'' are A-associated, then for all i,

$$U_i(A,s'_{-i}) - U_i(B,s'_{-i}) \tag{25}$$

$$
\begin{aligned}
&\underbrace{=}_{\text{by (6)}} \sum_{\substack{j\in N\setminus\{i\}\\ s'_j=A}} w_{ij}\,(a_{AA}-a_{BA}) - \sum_{\substack{j\in N\setminus\{i\}\\ s'_j=B}} w_{ij}\,(a_{BB}-a_{AB})\\
&\underbrace{\geq}_{\text{by (7)}} \sum_{\substack{j\in N\setminus\{i\}\\ s'_j=A}} w_{ij}\,(a_{BB}-a_{AB}) - \sum_{\substack{j\in N\setminus\{i\}\\ s'_j=B}} w_{ij}\,(a_{AA}-a_{BA})\\
&\underbrace{\geq}_{\substack{\text{by } A\text{-association}\\ \text{of } s', s''}} \sum_{\substack{j\in N\setminus\{i\}\\ s''_j=B}} w_{ij}\,(a_{BB}-a_{AB}) - \sum_{\substack{j\in N\setminus\{i\}\\ s''_j=A}} w_{ij}\,(a_{AA}-a_{BA})\\
&\underbrace{=}_{\text{by (6)}} U_i(B,s''_{-i}) - U_i(A,s''_{-i}).
\end{aligned}
$$

That is, (24) holds and s^A is GR-dominant. If risk dominance is strict, then the first $\geq$ in (25) is strict, therefore (24) holds strictly and s^A is strictly GR-dominant.

Step 2

The definition of risk dominance implies that at least one of A, B is risk dominant.

Non-strict

By Step 1, if A is risk dominant, then s^A is GR-dominant. Theorem 1 then implies that s^A is stochastically stable.

If A is not risk dominant, then B is strictly risk dominant. Step 1 then implies that s^B is strictly GR-dominant and Theorem 1 implies that s^B is uniquely stochastically stable. Therefore, s^A is not stochastically stable.

Strict

By Step 1, if A is strictly risk dominant, then s^A is strictly GR-dominant. Theorem 1 then implies that s^A is uniquely stochastically stable.

If A is not strictly risk dominant, then B is risk dominant. Step 1 then implies that s^B is GR-dominant and Theorem 1 implies that s^B is stochastically stable. Therefore, s^A is not uniquely stochastically stable. □

Proof of Proposition 2.

Risk dominance of A and Step 1 of the proof of Proposition 1 together imply that s^A is GR-dominant. Theorem 2 then implies that s^A is stochastically stable

Let $s' = (B, s^A_{-i})$. Let s'' be such that $s'_i \neq s''_i$ for all $i \in N$. Note that $B^k(s') = s''$ for odd k and $B^k(s') = s'$ for even k.

From s^A, player i switches to B with positive probability. Consequently, $c(s^A, s') = c_i(s^A, s') = 0$. It follows from the tree characterization of stochastically stable states that s' is also stochastically stable.

***k* odd**

Let s^* be any strategy profile. From s', with positive probability (that does not approach zero as $\eta \to 0$), all players in the set $\{j \neq i : s^*_j = B\}$ update their strategies, best responding to $B^{k-1}(s') = s'$ by playing B. Denote the resulting profile s^{**}. Note that $c(s', s^{**}) = 0$.

If $s^* = s^{**}$, then it follows from the tree characterization of stochastically stable states that s^* is stochastically stable.

If $s^* \neq s^{**}$, then it must be that $s_i^* \neq s_i^{**} = B$. From s^{**}, with positive probability (that does not approach zero as $\eta \to 0$) player i updates his strategy and switches to A. Therefore, $c(s^{**}, s^*) = 0$. It follows from the tree characterization of stochastically stable states that s^* is stochastically stable.

k even

Consider a profile s. Note that $B_i^1(s)$ is independent of s_i. If $B_i^1(s) = A$ (respectively, B), then for $j \neq i$, $B_j^2(s) = A$ (respectively, B), $B_i^3(s) = A$ (respectively, B), and so on. In particular, $B_i^{k-1}(s) = B_i^1(s)$ is independent of s_i. Therefore, the choice probabilities of $j \neq i$ are independent of s_i.

Now consider that if i were of level $k \geq 1$, Proposition 1 would imply that s^A is the unique stochastically stable profile. As the behavior of $j \neq i$ is independent of s_i, it must be that at any stochastically stable profile s^* we have $s_j^* = s_j^A = A$. Therefore no profile other than s^A and s' is stochastically stable. □

A.4 Proof of Theorem 3

Applying (15) to (9), we have

$$c_i(s,s') := \begin{cases} 0 & \text{if } s' = s, \\ g_i\left(\Upsilon_i(s_i', f_i(s))\right) & \text{if } s_i' \neq s_i,\ s_{-i}' = s_{-i}, \\ \infty & \text{otherwise .} \end{cases} \tag{26}$$

Lemma 4. *Let f preserve A-association.* ***(i)*** *If s^A is GR-dominant, then, for all $i \in N$, c_i is asymmetric towards A.* ***(ii)*** *If s^A is strictly GR-dominant and, for all $i \in N$, g_i is strictly increasing and $f_i(s^A) = s^A$, then, for all $i \in N$, c_i is strictly asymmetric towards A.*

Proof. Consider $s, s', \tilde{s} \in S$ such that $s, \tilde{s}$ are A-associated. As f preserves A-association, we have

$$f_i(s), f_i(\tilde{s}) \quad \text{are } A\text{-associated.} \tag{27}$$

Case 1: $s = s'$ or $s_i' = A$ or $c_i(s,s') = \infty$ or $\tilde{s}_i = A$.

If $c_i(s,s') = \infty$, let $\tilde{s}' = s^A$. The conditions in Definitions 4 and 5b are satisfied.

If $c_i(s,s')$ is finite, let $\tilde{s}' = \tilde{s}$. (26) implies $c_i(\tilde{s}, \tilde{s}') = 0$, therefore the conditions in Definitions 4 and 5b are satisfied.

Case 2: $s \neq s'$ and $s_i' \neq A$ and $c_i(s,s')$ is finite and $\tilde{s}_i \neq A$.

(26) and finiteness of $c_i(s,s')$ implies $s_{-i} = s_{-i}'$. $s \neq s'$ and $s_i' \neq A$ then imply that $s_i = A$. Let $\tilde{s}'$ be such that $\tilde{s}_{-i}' = \tilde{s}_{-i}$, $\tilde{s}_i' = A$.

If $c_i(\tilde{s}, \tilde{s}') = 0$, then the conditions in Definitions 4 and 5b are satisfied.

If $c_i(\tilde{s},\tilde{s}') > 0$, then

$$(28) \qquad \begin{aligned} c_i(s,s') &\underbrace{=}_{\text{by (26)}} g_i\left(\Upsilon_i(s_i', f_i(s))\right) \\ &\underbrace{=}_{\text{by (8)}} g_i\left(\max_{x_i\in S_i} U_i(x_i,(f_i(s))_{-i}) - U_i(s_i',(f_i(s))_{-i})\right) \\ &\geq g_i\left(U_i(A,(f_i(s))_{-i}) - \max_{x_i\neq A} U_i(x_i,(f_i(s))_{-i})\right) \\ &\underbrace{\geq}_{\substack{\text{by (27)}\\ \text{and GR-dominance}}} g_i\left(\max_{x_i\neq A} U_i(x_i,(f_i(\tilde{s}))_{-i}) - U_i(A,(f_i(\tilde{s}))_{-i})\right) \\ &\underbrace{=}_{\text{by (8)}} g_i\left(\Upsilon_i(\tilde{s}_i', f_i(\tilde{s}))\right) \underbrace{=}_{\substack{\text{by (26)}\\ \text{and } c(\tilde{s},\tilde{s}')>0}} c(\tilde{s},\tilde{s}'). \end{aligned}$$

That is, the condition in Definition 4 as satisfied. Further note that if s^A is strictly GR-dominant and g_i is strictly increasing, then the weak inequality in (28) due to GR-dominance becomes a strict inequality, so that the condition in Definition 5b is satisfied.

For the remainder of this proof, assume that $f(s^A) = s^A$, s^A is strictly GR-dominant, and g_i is strictly increasing.

Note that, as s^A is A-associated with itself, strict GR-dominance of s^A implies that

$$(29) \qquad \arg\max_{x_i\in S_i} U_i(x_i,(s^A)_{-i}) = \{A\}.$$

If $s \neq s^A$, then either $s_{-i} \neq s^A_{-i}$, in which case $c(s^A,s) = \infty$, or $s_{-i} = s^A_{-i}$, $s_i \neq A$, in which case

$$(30) \qquad \begin{aligned} c(s^A,s) &\underbrace{=}_{\text{by (26)}} g_i\left(\Upsilon_i(s_i, f_i(s^A))\right) \\ &\underbrace{=}_{\text{by (8)}} g_i\left(\max_{x_i\in S_i} U_i(x_i,(f_i(s^A))_{-i}) - U_i(s_i,(f_i(s^A))_{-i})\right) \\ &\underbrace{=}_{\text{by } f(s^A)=s^A} g_i\left(\max_{x_i\in S_i} U_i(x_i,(s^A)_{-i}) - U_i(s_i,(s^A)_{-i})\right) \\ &\underbrace{>}_{\substack{\text{by } s_i\neq A, (29),\\ g_i \text{ strictly increasing}}} 0. \end{aligned}$$

Therefore, the condition in Definition 5a is satisfied. □

Proof of Theorem 3.

Assume s^A is GR-dominant. Lemma 4(i) implies that, for all $i \in N$, c_i is asymmetric towards A.

Theorem N(i) then implies that c is asymmetric towards A. Theorem P(i) then implies that s^A is stochastically stable.

Assume s^A is strictly GR-dominant and, for all $i \in N$, g_i is strictly increasing and $f_i(s^A) = s^A$. Lemma 4(ii) implies that, for all $i \in N$, c_i is strictly asymmetric towards A. Theorem N(ii) then implies that c is strictly asymmetric towards A. Theorem P(ii) then implies that s^A is uniquely stochastically stable. □

A.5 Proof of Theorem 4

Let P^{ε} be the Markov kernel of the process with $F = \{f_1, \dots, f_n\}$. Define processes $P^{\varepsilon, f_1}, \dots P^{\varepsilon, f_n}$ as identical to P^{ε} except that $F = \{f_1\}, \dots, F = \{f_n\}$ respectively. Note that

$$P^{\varepsilon} = \sum_{i=1}^{n} \varphi(f_m) P^{\varepsilon, f_i}. \tag{31}$$

Let cost functions for $P^{\varepsilon, f_1}, \dots P^{\varepsilon, f_n}$ be given by $c^{f_1}, \dots, c^{f_n}$.

Proof of Theorem 4.

Assume s^A is GR-dominant. Lemma 4(i) implies that, for all $i \in N$, $f_m \in F$, $c_i^{f_m}$ is asymmetric towards A. Theorem N(i) then implies that, for $m = 1, \dots, n$, c^{f_m} is asymmetric towards A. Given (31), that is P^{ε} is a convex combination of $P^{\varepsilon, f_1}, \dots P^{\varepsilon, f_n}$, this implies that c is asymmetric towards A (Newton, 2020, Theorem 1). Theorem P(i) then implies that s^A is stochastically stable.

Assume s^A is strictly GR-dominant and, for all $i \in N$, g_i is strictly increasing and $f_i(s^A) = s^A$. Lemma 4(ii) implies that, for all $i \in N$, $f_m \in F$, $c_i^{f_m}$ is strictly asymmetric towards A. Theorem N(ii) then implies that, for $m = 1, \dots, n$, c^{f_m} is strictly asymmetric towards A. Given (31), that is P^{ε} is a convex combination of $P^{\varepsilon, f_1}, \dots P^{\varepsilon, f_n}$, this implies that c is strictly asymmetric towards A (Newton, 2019, Theorem 1). Theorem P(ii) then implies that s^A is uniquely stochastically stable.

□

A.6 Proofs for Section 7

Proof of Theorem 5.

Non-strict

If A is a weakly dominant strategy for i, then

$$U_i(A, s_{-i}) \geq \max_{x_i \neq A} U_i(x_i, s_{-i}) \quad \text{for all } s_{-i}, \tag{32}$$

and consequently

$$U_i(A, s'_{-i}) + U_i(A, s''_{-i}) \geq \max_{s_i \neq A} U_i(s_i, s'_{-i}) + \max_{s_i \neq A} U_i(s_i, s''_{-i}). \tag{33}$$

Therefore, if A is weakly dominant for all $i \in N$, then (33) holds for all i, s', s''. Therefore, s^A is GR-dominant.

Strict

If A is a strictly dominant strategy for i, then

$$(34) \qquad U_i(A, s'_{-i}) > \max_{x_i \neq A} U_i(x_i, s_{-i}) \quad \text{for all } s_{-i},$$

and consequently

$$(35) \qquad U_i(A, s'_{-i}) + U_i(A, s''_{-i}) > \max_{s_i \neq A} U_i(s_i, s'_{-i}) + \max_{s_i \neq A} U_i(s_i, s''_{-i}).$$

Therefore, if A is strictly dominant for all $i \in N$, then (35) holds for all i, s', s''. Therefore, s^A is strictly GR-dominant. □

Proof of Proposition 3.

For all $i \in N$, it is a weakly dominant strategy for player i to bid his valuation. Therefore, s^A is composed of weakly dominant strategies and Theorem 5 implies that s^A is GR-dominant. Theorem 4 then implies that s^A is stochastically stable. □

Proof of Proposition 4.

Let s' and s'' be A-associated.

Consider $s^*_i \neq A$. There must exist $v^*_i \in V_i$ such that $s^*_i(v^*_i) \neq v^*_i$. The restriction to sensible strategies implies that $s^*_i(v^*_i) \leq v^*_i$, so it must be that $s^*_i(v^*_i) < v^*_i$. Note that $s^*_i(v^*_i) \in B_i = V_i$.

As s' and s'' are A-associated, for a given player $j \neq i$, it must be that either $s'_j = A$ or $s''_j = A$. Without loss of generality assume that $s'_j = A$.

<u>Case 1</u>: Player i has valuation v^*_i and all other players $k \neq i$ have valuation $v^*_k = s^*_i(v^*_i)$.

As $s'_j = A$, we have $s'_j(v^*_j) = s'_j(s^*_i(v^*_i)) = s^*_i(v^*_i)$.

Sensible strategies imply that for $k \neq i, j$, we have $s'_k(v^*_k) = s'_k(s^*_i(v^*_i)) \leq s^*_i(v^*_i)$.

Consider strategy profile (s^*_i, s'_{-i}). Player i's bid is $s^*_i(v^*_i)$. This equals player j's bid and is greater than or equal to all other bids. Therefore, player i wins the good and obtains a realized payoff of $v^*_i - s^*_i(v^*_i) > 0$ with probability at most one half.

Consider strategy profile (A, s'_{-i}). Player i's bid is v^*_i. This is strictly greater than all other bids. The second highest bid (possibly tied) is player j's bid of $s^*_i(v^*_i)$. Therefore, player i wins the good and obtains a realized payoff of $v^*_i - s^*_i(v^*_i) > 0$ with probability one.

Therefore at valuations $(v^*_k)_{k \in N}$, player i obtains a strictly higher realized payoff from (A, s'_{-i}) than he does from (s^*_i, s'_{-i}).

<u>Case 2</u>: Consider all valuations $(v_k)_{k \in N}$ other than $(v^*_k)_{k \in N}$ considered in Case 1. By standard arguments for second price auctions, at any realized valuations, a player with strategy A could not have obtained a strictly higher payoff by playing a strategy other than A.

Therefore at valuations $(v_k)_{k \in N}$, player i obtains at least as high a payoff from (A, s'_{-i}) as he does from (s^*_i, s'_{-i}).

As v_k are iid with full support on $V_k = V_i$, Case 1 occurs with positive probability. Therefore, combining Case 1 and Case 2, we obtain

$$U_i(A, s'_{-i}) > U_i(s^*_i, s'_{-i}). \tag{36}$$

Next, note that weak dominance of A implies that

$$U_i(A, s''_{-i}) \geq U_i(s^*_i, s''_{-i}). \tag{37}$$

Summing (36) and (37),

$$U_i(A, s'_{-i}) + U_i(A, s''_{-i}) > U_i(s^*_i, s'_{-i}) + U_i(s^*_i, s''_{-i}). \tag{38}$$

Our argument and hence (38) holds for any $i \in N$ and A-associated s' and s''. That is, s^A is strictly GR-dominant according to Definition 1.

Applying Theorem 4, it follows that s^A is uniquely stochastically stable. □

References

Alós-Ferrer, C. and Netzer, N. (2010). The logit-response dynamics. *Games and Economic Behavior*, 68:413–427.

Bacharach, M. (1992). Backward induction and beliefs about oneself. *Synthese*, 91(3):247–284.

Battalio, R., Samuelson, L., and Van Huyck, J. (2001). Optimization incentives and coordination failure in laboratory stag hunt games. *Econometrica*, 69(3):749–764.

Block, H. and Marschak, J. (1959). Random orderings and stochastic theories of response. Working paper, Cowles Foundation for Research in Economics, Yale University.

Blume, L. E. (1993). The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5(3):387 – 424.

Blume, L. E. (1996). Population games. Working Papers 96-04-022, Santa Fe Institute.

Blume, L. E. (2003). How noise matters. *Games and Economic Behavior*, 44(2):251–271.

Cabrales, A., Nagel, R., and Armenter, R. (2007). Equilibrium selection through incomplete information in coordination games: an experimental study. *Experimental Economics*, 10(3):221–234.

Conlisk, J. (1993a). Adaptation in games: Two solutions to the crawford puzzle. *Journal of Economic Behavior & Organization*, 22(1):25–50.

Conlisk, J. (1993b). Adaptive tactics in games: Further solutions to the crawford puzzle. *Journal of Economic Behavior & Organization*, 22(1):51–68.

Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 215–242.

Crawford, V. (2007). Modeling behavior in novel strategic situations via level-k thinking. Conference slides - North American Economic Science Association Meetings.

Crawford, V. P. (1974). Learning the optimal strategy in a zero-sum game. *Econometrica: Journal of the Econometric Society*, pages 885–891.

Crawford, V. P. (2019). Experiments on cognition, communication, coordination, and cooperation in relationships.

Duffy, J. and Nagel, R. (1997). On the robustness of behaviour in experimental 'beauty contest'games. *The Economic Journal*, 107(445):1684–1700.

Foster, D. and Young, H. P. (1990). Stochastic evolutionary game dynamics. *Theoretical Population Biology*, 38:219–232.

Geanakoplos, J. and Gray, L. (1991). When seeing further is not seeing better. *Bulletin of the Santa Fe Institute*, 6(2):1–6.

Gill, D. and Prowse, V. (2016). Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. *Journal of Political Economy*, 124(6):1619–1676.

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

Harsanyi, J. C. and Selten, R. (1988). A general theory of equilibrium selection in games. *MIT Press Books*, 1.

Heinemann, F., Nagel, R., and Ockenfels, P. (2004). The theory of global games on test: experimental analysis of coordination games with public and private information. *Econometrica*, 72(5):1583–1599.

Heller, Y. (2015). Three steps ahead. *Theoretical Economics*, 10(1):203–241.

Ho, T.-H. and Su, X. (2013). A dynamic level-k model in sequential games. *Management Science*, 59(2):452–469.

Hwang, S.-H., Lim, W., Neary, P., and Newton, J. (2018). Conventional contracts, intentional behavior and logit choice: Equality without symmetry. *Games and Economic Behavior*, 110:273 – 294.

Iijima, R. (2015). Iterated generalized half-dominance and global game selection. *Journal of Economic Theory*, 159:120–136.

Kandori, M., Mailath, G. J., and Rob, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56.

Khan, A. and Peeters, R. (2014). Cognitive hierarchies in adaptive play. *International Journal of Game Theory*, 43(4):903–924.

Lim, W. and Neary, P. R. (2016). An experimental investigation of stochastic adjustment dynamics. *Games and Economic Behavior*, 100:208–219.

Mäs, M. and Nax, H. H. (2016). A behavioral study of "noise" in coordination games. *Journal of Economic Theory*, 162:195–208.

Matros, A. (2003). Clever agents in adaptive learning. *Journal of Economic Theory*, 111(1):110–124.

Mohlin, E. (2012). Evolution of theories of mind. *Games and Economic Behavior*, 75(1):299–318.

Morris, S., Rob, R., and Shin, H. S. (1995). p-dominance and belief potential. *Econometrica*, 63:145–157.

Myatt, D. P. and Wallace, C. (2003). Sophisticated play by idiosyncratic agents. *Journal of Evolutionary Economics*, 13(3):319–345.

Myerson, R. (1978). Refinements of the Nash equilibrium concept. *International Journal of Game Theory*, 7:73–80.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5):1313–1326.

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49.

Newton, J. (2017). Shared intentions: The evolution of collaboration. *Games and Economic Behavior*, 104:517 – 534.

Newton, J. (2018). Evolutionary game theory: A renaissance. *Games*, 9(2):31.

Newton, J. (2019). Combining strictly asymmetric choice rules. SSRN Working Paper Series 3433736.

Newton, J. (2020). Conventions under heterogeneous behavioural rules. *Review of Economic Studies*, (https://doi.org/10.1093/restud/rdaa063).

Peski, M. (2010). Generalized risk-dominance and asymmetric dynamics. *Journal of Economic Theory*, 145(1):216 – 248.

Proto, E., Rustichini, A., and Sofianos, A. (2019). Intelligence, personality, and gains from cooperation in repeated interactions. *Journal of Political Economy*, 127(3):1351–1390.

Robalino, N. and Robson, A. (2016). The evolution of strategic sophistication. *American Economic Review*, 106(4):1046–1072.

Rusch, H. (2019). The evolution of collaboration in symmetric 2×2-games with imperfect recognition of types. *Games and Economic Behavior*.

Sáez-Martı, M. and Weibull, J. W. (1999). Clever agents in young's evolutionary bargaining model. *Journal of Economic Theory*, 86(2):268–279.

Sandholm, W. H. (2010). *Population games and evolutionary dynamics*. Economic learning and social evolution. Cambridge, Mass. MIT Press.

Selten, R. (1991). Anticipatory learning in two-person games. In *Game equilibrium models I*, pages 98–154. Springer.

Stahl, D. O. (1993). Evolution of smartn 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(1):105–138.

Stahl, D. O. and Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior and Organization*, 25(3):309–327.

Stennek, J. (2000). The survival value of assuming others to be rational. *International Journal of Game Theory*, 29(2):147–163.

Tang, F.-F. (2001). Anticipatory learning in two-person games: some experimental results. *Journal of economic behavior & organization*, 44(2):221–232.

Van Huyck, J. B., Battalio, R. C., and Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review*, pages 234–248.

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

Young, H. P. (1993b). An evolutionary model of bargaining. *Journal of Economic Theory*, 59(1):145 – 168.