No. 28 june 2008

Carlos Alós-Ferrer
Nick Netzer

The Logit-Response Dynamics

Research Paper Series
Thurgau Institute of Economics and Department of Economics
at the University of Konstanz

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

The Logit-Response Dynamics*

Carlos Alós–Ferrer† and Nick Netzer‡

This Version: June 2008

First Version: October 2007

Abstract

We develop a characterization of stochastically stable states for the logit-response learning dynamics in games, with arbitrary specification of revision opportunities. The result allows us to show convergence to the set of Nash equilibria in the class of best-response potential games and the failure of the dynamics to select potential maximizers beyond the class of exact potential games. We also study to which extent equilibrium selection is robust to the specification of revision opportunities. Our techniques can be extended and applied to a wide class of learning dynamics in games.

Keywords: Learning in games, logit-response dynamics, best-response potential games.

JEL Classification Numbers: C72, D83.

---

*We thank Hans Haller, Josef Hofbauer, Arthur Robson and Bill Sandholm for helpful comments. Financial support from the Austrian Science Fund (FWF) under Project P18141-G09 is gratefully acknowledged.

†University of Konstanz, Department of Economics, Box 150, D-78457 Konstanz, Germany. E-mail: Carlos.Alos-Ferrer@uni-konstanz.de

‡University of Zurich, Socioeconomic Institute, Bluemlisalpstrasse 10, CH-8006 Zurich, Switzerland. E-mail: Nick.Netzer@soi.uzh.ch.

1 Introduction

Models of learning in games typically start with the specification of a basic behavioral rule on the side of the players, e.g. myopic best reply, truncated fictitious play, or a variant of imitation. Since such basic dynamics exhibit a multiplicity of rest points (e.g., any Nash equilibrium is a rest point for a myopic best reply dynamics), it is necessary to perform a stability test.

Within the class of discrete-time, finite population models, one of the most successful paradigms in the literature performs this test by adding *noise* to the basic dynamics and studying the long-run outcomes as noise vanishes. Formally, the basic dynamics is a Markov chain[1] with multiple absorbing sets, which is made irreducible by the addition of noise. Probably the best-known example of this methodology is the *mistakes model*, essentially introduced by Kandori, Mailath, and Rob (1993) (for an imitation rule), Young (1993), and Kandori and Rob (1995) (for myopic best reply). In this model, agents are assumed to have a certain probability (independent across agents and periods) of making mistakes, where a mistake is defined as choosing some strategy at random, with a full-support probability distribution. The mistake distribution is typically assumed to be uniform, although this is of no relevance. The important feature of the model is that the shape of this distribution is independent of the noise level.

The mistakes model has delivered important messages, ranging from the almost universal selection of risk-dominant equilibria (as opposed to Pareto efficient ones) in coordination games (Kandori, Mailath, and Rob (1993), Young (1993)), to the dynamic relevance of "perfectly competitive" outcomes in aggregative games (Vega-Redondo (1997), Alós-Ferrer and Ania (2005)). One of the most attractive features of the mistakes model is that, thanks to a result due to Freidlin and Wentzell (1988), it is possible to provide a simple characterization of the set of long-run outcomes. These outcomes, called *stochastically stable states*, are those having positive probability in the limit

[1]Throughout this paper, the term "Markov chain" refers to a discrete-time Markov chain with stationary transition probabilities and finite state space.

invariant distribution as noise vanishes. The well-known characterization relies on minimizing the number of mutations associated to the transitions depicted in certain graphs (trees) defined on the state space.

This approach is not exempt of critiques. First and foremost, selection results are based on the number of mistakes necessary to destabilize a given state, but, in a sense, all mistakes are treated equally. For example, choosing a strategy which delivers *almost* a best response is just as severe a mistake as choosing a strategy which delivers payoffs far away from the optimum. Thus, the approach relies on the (cardinal) payoffs of the game only to a limited extent. Second, Bergin and Lipman (1996) observed that, if the distribution of mistakes is allowed to be state-dependent, the model can be twisted to select any pre-specified rest point. Thus, it becomes necessary to have a well-justified theory on the origin of mistakes.[2] One particular model that accounts for both problems is the *logit response* dynamics, which was introduced in Blume (1993). In this dynamics, which can be derived from a random utility model, players adopt an action according to a full-support distribution of the logit form, which allocates larger probability to those actions which would deliver (myopically) larger payoffs. It therefore combines the advantage of having a specific theory about the origin of mistakes with the fact that it takes the magnitude of (suboptimal) payoffs fully into account. Noise is incorporated in the specification from the onset, but choices concentrate on best responses as noise vanishes.

The logit-response dynamics is not a particular case of the mistakes model, and thus cannot benefit from the characterization of long-run outcomes mentioned above. Indeed, results for the logit-response dynamics are harder to obtain and are restricted to particularly well-behaved classes of games. For example, binary action games (as in Blume (2003) or Maruta (2002)) give rise to a birth-death chain whose invariant distribution can be

[2]In van Damme and Weibull (2002), mistakes arise in a rationalistic model where agents have to exert costly effort to reduce mistake probabilities. Blume (2003) responds to the Bergin and Lipman critique by characterizing the class of noise processes that select risk-dominant equilibria in coordination games.

characterized directly. Further, as shown by Blume (1997), if the base game admits an exact potential, the process is reversible and again the invariant distribution can be characterized directly. In addition to the restricted class of games, the existing results also rely on specific assumptions about revision opportunities, most notably on the assumption of one-at-a-time updating (asynchronous learning). Given the sensitivity of the original mistakes model to the specification of the dynamics, where it is sometimes the case that "who learns when" is as important as "who learns how", studying robustness issues is fundamental.

In this paper we develop a characterization of the long-run outcomes of the logit-response dynamics for arbitrary finite normal-form games. Furthermore, our result applies to a generalization of the original logit-response dynamics. In particular, we allow for an arbitrary specification of revision opportunities, encompassing e.g. independent revision opportunities (as in most versions of the mistakes model) and asynchronous learning (as in Blume's (1993) model). In order to obtain our results, we build on the analysis of Freidlin and Wentzell (1988) to characterize the invariant distribution of the logit-response dynamics for fixed noise levels, and use it to develop a characterization of the stochastically stable states.

In order to illustrate the method and its applicability, we proceed to study the convergence of the logit-response model for the various generalizations of the concept of potential game. Our method allows us to offer simple answers to several open questions. We find that, first, convergence to the set of Nash equilibria cannot be guaranteed for Monderer and Shapley's (1996) generalized ordinal potential games, but, second, convergence does obtain for Voorneveld's (2000) best-response potential games. We also show that the latter result is robust to the specification of revision opportunities under an additional condition which is satisfied both by independent inertia and asynchronous learning models. Third, we study the value of potential maximizers as an equilibrium refinement and find that the selection of potential maximizers (which obtains for exact potential games under Blume's

(1993) asynchronous learning dynamics) fails two robustness tests. First, it fails even for exact potential games if revision opportunities do not fall into the asynchronous learning category. Second, it fails for any generalization of potential games even if revision opportunities are asynchronous.

The paper is organized as follows. Section 2 reviews the logit-response dynamics and introduces our generalized dynamics. Section 3 presents our characterization of stochastically stable states, whose (technical) proof is relegated to the Appendix. Section 4 applies this characterization to the logit-response dynamics in best-response potential games. Section 5 discusses a generalization of our results. Section 6 concludes.

2 The Logit-Response Dynamics

2.1 The Logit Choice Function

Let $\Gamma = (I, (S_i, u_i)_{i \in I})$ be a finite normal-form game with player set $I = \{1, 2, ..., N\}$, strategy sets S_i and payoff functions u_i defined on the set of pure strategy profiles $S = S_1 \times ... \times S_N$. For a given player i, denote by $S_{-i} = \prod_{j \neq i} S_j$ the set of pure strategy profiles of i's opponents. Following convention, we denote $s = (s_i, s_{-i}) \in S$ and $u_i(s_i, s_{-i}) = u_i(s)$.

The game is played by boundedly rational players, who behave as myopic best repliers, but tremble in their decisions. Every period, some set of players is chosen to update their actions. We will further specify revision opportunities below.

When given the chance to revise, player i observes the actions s_{-i} of the opponents. The probability of choosing action s_i given the current profile s_{-i} is given by the *logit choice function*

$$p_i(s_i, s_{-i}) = \frac{e^{\beta u_i(s_i, s_{-i})}}{\sum_{s'_i \in S_i} e^{\beta u_i(s'_i, s_{-i})}}, \tag{1}$$

where $0 < \beta < \infty$.

The scalar β can be interpreted as an inverse measure of the level of noise in players' decisions. As $\beta \to \infty$, the described rule converges to the myopic best reply rule. For any $0 < \beta < \infty$, players choose non-best replies with positive probability, but actions that yield smaller payoffs are chosen with smaller probability. The dynamic adjustment process defines an irreducible and aperiodic Markov chain $\{X_t^\beta\}_{t\in\mathbb{N}}$ on the state space S, with stationary transition probabilities $P_{ss'}^\beta = \Pr^\beta(s_t = s'|s_{t-1} = s)$ and (unique) invariant distribution μ^β.

For any specification of revision opportunities, we will refer to this dynamics as a *logit response dynamics*. Consider the particular case where exactly one player is randomly selected each period to revise his strategy,[3] and let $q_i > 0$ denote the probability that player i is selected. For this case, which we will refer to as *asynchronous learning*, the dynamics was first introduced (with $q_i = \frac{1}{/}N$) by Blume (1993)[4] and has been further developed in e.g. Blume (1997, 2003), Young (1998), and Baron, Durieu, Haller, and Solal (2002a, 2002b). Taken as a behavioral rule, the underlying logit choice function (1) is rooted in the psychology literature (Thurstone (1927)). From the microeconomic point of view, it can be given a justification in terms of a random-utility model (see e.g. McKelvey and Palfrey (1995) for details). Hofbauer and Sandholm (2002, Section 2) observe that it is also the only choice function of the "quantal" form

$$C_i(u_i) = \frac{w(u_i(s_i, s_{-i}))}{\sum_{s_i' \in S_i} w(u_i(s_i', s_{-i}))}$$

with $w(\cdot)$ an increasing and differentiable function of the payoffs, which can be derived as the result of both a stochastic and a deterministic perturbation of the payoffs.[5] Thus, the logit-response dynamics exhibits solid decision-theoretic foundations.

[3]This can be interpreted as a reduced form (technically, the *embedded chain*) of a continuous-time model where players receive revision opportunities according to "Poisson alarm clocks."

[4]Blume (1993) refers to this dynamics as log-linear response.

[5]Mattsson and Weibull (2002) and Baron, Durieu, Haller, and Solal (2002a, 2002b) show that logit-response arises in the framework of van Damme and Weibull (2002) when control costs adopt a specific functional form.

2.2 Asynchronous Logit Response in Potential Games

The game Γ is a *potential game*[6] (Monderer and Shapley (1996)) if there exists a function $\rho : S \to \mathbb{R}$, called the potential, such that for each $i \in I$, $s_i, s_i' \in S_i$, $s_{-i} \in S_{-i}$

$$u_i(s_i, s_{-i}) - u_i(s_i', s_{-i}) = \rho(s_i, s_{-i}) - \rho(s_i', s_{-i}).$$

The global maximizers of the potential function ρ form a subset of the set of Nash equilibria of Γ. If Γ is a potential game, it follows that $u_i(s_i, s_{-i}) = \rho(s_i, s_{-i}) + \lambda(s_{-i})$, where $\lambda(s_{-i})$ is independent of s_i.[7] Thus (1) can be simplified to

$$p_i(s_i, s_{-i}) = \frac{e^{\beta u_i(s_i, s_{-i})}}{\sum_{s_i' \in S_i} e^{\beta u_i(s_i', s_{-i})}} = \frac{e^{\beta \rho(s_i, s_{-i})}}{\sum_{s_i' \in S_i} e^{\beta \rho(s_i', s_{-i})}}. \tag{2}$$

It is then straightforward to show (see Blume (1997)) that the invariant distribution of the logit-response dynamics adopts a *Gibbs-Boltzmann* form, i.e. the potential function becomes a potential for the stochastic process. The proof (which is included for completeness only) takes advantage of the fact that the reformulation (2) implies that the process is reversible.

Proposition 1. *Let Γ be a potential game with potential ρ. The invariant distribution of the logit-response dynamics with asynchronous learning is*

$$\mu^\beta(s) = \frac{e^{\beta \rho(s)}}{\sum_{s' \in S} e^{\beta \rho(s')}}.$$

Proof. It is enough to show that μ^β as given in the statement satisfies the *detailed balance condition*, i.e. $\mu^\beta(s) P^\beta_{ss'} = \mu^\beta(s') P^\beta_{s's}$ for all $s, s' \in S$. This is clearly fulfilled if $s = s'$, and also if s and s' differ in more than one coordinate, since $P^\beta_{ss'} = P^\beta_{s's} = 0$ in this case. Hence assume w.l.o.g. that s and s' differ exactly in coordinate i, that is $s_i \neq s_i'$ and $s_j = s_j'$ for all $j \neq i$. It follows that

$$\mu^\beta(s) P^\beta_{ss'} = \frac{e^{\beta \rho(s)}}{\sum_{s'' \in S} e^{\beta \rho(s'')}} \, q_i \, \frac{e^{\beta \rho(s_i', s_{-i})}}{\sum_{s_i'' \in S_i} e^{\beta \rho(s_i'', s_{-i})}} = \mu^\beta(s') P^\beta_{s's}.$$

[6] Also called *partnership games*. See Hofbauer and Sigmund (1988).

[7] Fix a strategy $s_0 \in S_i$, and define $\lambda(s_{-i}) = u_i(s_0, s_{-i}) - \rho(s_0, s_{-i})$.

where the last equality holds due to $s_{-i} = s'_{-i}$. ■

As $\beta \to \infty$, the invariant distribution of the process converges to an invariant distribution of the best-reply dynamics. We say that a state $s \in S$ is *stochastically stable* if $\lim_{\beta\to\infty} \mu^\beta(s) > 0$. An immediate consequence of Proposition 1 is

Corollary 1. *Let Γ be a potential game. The set of stochastically stable states of the logit-response dynamics with asynchronous learning is equal to the set of maximizers of ρ.*

This Corollary provides of course a readily applicable result.[8] In our view, it is also important for two additional reasons. First, it is a *convergence* result. The asynchronous logit response dynamics always converges to the set of Nash equilibria in the class of exact potential games. Second, it is a *selection* result. In particular, the logit-response dynamics provides support for treating the set of potential maximizers as an equilibrium refinement for potential games.

The latter finding has motivated a large part of the literature of learning in games in recent years, and indeed the selection of potential maximizers has become a test of the reasonability of a learning dynamics.[9] It is therefore important to know how robust both parts of Corollary 1 are. That is, we pose the question of whether the convergence to Nash equilibria and the selection of potential maximizers extend to more general classes of games and dynamics.

Proposition 1 (and hence Corollary 1), however, rely on the knife-edge technical fact that the exact potential of the game allows to identify the invariant distribution of the stochastic process for positive noise level. Clearly, the proof cannot be generalized any further. In the next section, we develop a framework which will allow us to provide exact results for both more general games and more general dynamics.

[8]For example, Sandholm (2007) relies on this result to build a model of evolutionary implementation.

[9]See e.g. Hofbauer and Sorger (1999).

2.3 Revision Processes and a Generalized Dynamics

The existing results for the logit-response rule (as e.g. Corollary 1) rely on the asynchronicity assumption to establish the convenient Gibbs-Boltzmann form for the invariant distribution. Here we will consider a more general approach allowing for arbitrary specification of updating opportunities. We illustrate this by considering a general class of revision processes. The motivation for the generalization is as follows. In our view, a learning dynamics in games is made of a behavioral rule and a specification of revision opportunities (i.e. the speed of the dynamics). Thus, it is important to know which results are due to the behavioral rule and which ones hinge on the exact specification of the revision process. Studying general revision processes for a given dynamics therefore becomes an important robustness check.

Definition 1. A *revision process* is a probability measure q on the set of subsets of I, $\mathcal{P}(I)$, such that

$$\forall\, i \in I\ \exists J \subseteq I \text{ such that } i \in J \text{ and } q_J > 0 \tag{3}$$

where, for each $J \subseteq I$, $q_J = q(J)$ is interpreted as the probability that *exactly* players in J receive revision opportunities (independently across periods).

Condition (3) merely specifies that every player has some probability of being able to revise in some situation. No further restriction is placed on the revision process, which allows for a wide range of models to be considered. We list now three leading examples.

Let $\mathcal{R}^q = \{J \subseteq I | q_J > 0\}$ denote the set of *revising sets*, i.e. sets of players which might obtain revision opportunity (as a whole) with positive probability. If $\mathcal{R}^q = \{\{i\}|i \in I\}$, we say that the dynamics exhibits *asynchronous learning* and write $q_i = q_{\{i\}}$. As commented above, this includes the asynchronous logit-response dynamics of Blume (1993) (where $q_i = \frac{1}{N}$).

If $\mathcal{R}^q = \mathcal{P}(I)$, we speak of *independent learning*. That is, every subset of players has positive probability of being able to revise. For example, a standard version of the mistakes model (see e.g. Sandholm (1998)) is a particular case which postulates *independent inertia*, i.e. each player revises with a fixed, independent probability $0 < p < 1$. Thus $q_J = p^{|J|}(1-p)^{N-|J|} > 0$ for each subset J.

We can also consider a model of *instantaneous learning*, where all players receive revision opportunities every period, i.e. $\mathcal{R}^q = \{I\}$. Other examples could include specific correlation in revision opportunities among certain groups of players,[10] or bounds to the number of players revising each period.

Fix a revision process q. For any two strategy profiles $s, s' \in S$, let $R_{s,s'} = \{J \in \mathcal{R}^q | s'_k = s_k \forall\ k \notin J\}$ be the set of revising sets potentially leading from s to s'. Note that from a given $s \in S$, different alternative revising sets might give rise to the same transition, because players selected to revise might stay with their previous action. However, under asynchronous learning $|R_{s,s'}| \leq 1$ for all $s \neq s'$. We say that a transition from s to s' is *feasible* if $R_{s,s'} \neq \emptyset$.

The logit-response dynamics with revision process q is a Markov chain on the state space S with stationary transition probabilities given by

$$P_{s,s'} = \sum_{J \in R_{s,s'}} q_J \prod_{j \in J} \frac{e^{\beta \cdot u_j(s'_j, s_{-j})}}{\sum_{s''_j \in S_j} e^{\beta \cdot u_j(s''_j, s_{-j})}}.$$

Define $U_J(s', s) = \sum_{j \in J} u_j(s'_j, s_{-j})$. Let $R^J_s = \{s' \in S | s'_k = s_k \forall\ k \notin J\}$ be the set of states potentially reached from s when the revising set is J. We can then rewrite the transition probabilities as

$$P_{s,s'} = \sum_{J \in R_{s,s'}} q_J \frac{e^{\beta \cdot U_J(s', s)}}{\sum_{s'' \in R^J_s} e^{\beta \cdot U_J(s'', s)}}. \tag{4}$$

3 Stochastic Stability

Given a general revision process q, the logit-response dynamics is in general not a birth-death chain. Even if this were the case (e.g. under asynchronicity in binary action games), unless the game is an (exact) potential game an exact characterization of the invariant distribution was until now not available. We

[10] Since we do not restrict attention to symmetric games, this possibility might be of independent interest, e.g. for buyers-sellers models as in Alós-Ferrer and Kirchsteiger (2007).

now develop a characterization of the set of stochastically stable states of the logit-response dynamics which relies precisely on such a characterization of the invariant distribution.

Given a state s, define an *s-tree* to be a directed graph T such that there exists a unique path from any state $s' \in S$ to s. The key concept for our characterization is as follows:

Definition 2. A *revision s-tree* is a pair (T, γ) where

(i) T is an s-tree,

(ii) $(s, s') \in T$ only if $R_{s,s'} \neq \emptyset$ (only feasible transitions are allowed), and

(iii) $\gamma : T \to \mathcal{P}(I)$ is such that $\gamma(s, s') \in R_{s,s'}$ for all $(s, s') \in T$.

Thus, there are two differences between a revision tree and a tree as used in the characterization for the mistakes model. First, in a revision s-tree, edges corresponding to unfeasible transitions are not allowed.[11] Second, in a revision s-tree (T, γ), γ labels each edge of T with a revising set which makes the corresponding transition potentially feasible.

Remark 1. Suppose that a revision process satisfies that for all $s, s' \in S$, $s \neq s'$, either $\mathcal{R}_{s,s'} = \emptyset$ or $|\mathcal{R}_{s,s'}| = 1$. This is e.g. true for asynchronous learning and instantaneous learning. Then, for each link (s, s') in a revision tree there exists exactly one revising set making the transition from s to s' feasible. In other words, given a tree T using feasible transitions only, there exists a unique mapping γ such that (T, γ) is a revision tree.

3.1 A Characterization

Let $\mathcal{T}(s)$ denote the set of revision s-trees. The *waste* of a revision tree $(T, \gamma) \in \mathcal{T}(s)$ is defined as

$$W(T, \gamma) = \sum_{(s,s') \in T} \left(\max_{s'' \in S} U_{\gamma(s,s')}(s'', s) \right) - U_{\gamma(s,s')}(s', s).$$

[11] Thus, actually the concept of revision tree depends on the revision process q. We drop this dependency for notational simplicity.

or, equivalently,

$$W(T,\gamma) = \sum_{(s,s')\in T} \sum_{j\in\gamma(s,s')} \left(\max_{s''_j \in S_j} u_j(s''_j, s_{-j}) - u_j(s'_j, s_{-j}) \right).$$

In words, the waste of a revision tree adds all the individual (ex-ante, myopic) payoff wastes generated across the transitions depicted in the tree, relative to the payoffs that could have been reached by adopting best responses. Obviously, a transition generates zero waste in this sum if and only if it involves only best responses.

Intuitively, the waste of a revision tree is an inverse measure for its likelihood in the logit-response dynamics. It is analogous to the concept of costs in the mistakes model, with the obvious difference that wastes are real numbers, rather than natural ones (number of mistakes).[12] The *stochastic potential* of a given state is obtained by minimizing waste across revision trees rooted in that state.

Definition 3. The *stochastic potential* of a state s is

$$W(s) = \min_{(T,\gamma)\in\mathcal{T}(s)} W(T,\gamma).$$

As mentioned above, a state is stochastically stable if it has positive probability in the limit invariant distribution of a noisy process as noise vanishes (in our case, when $\beta \to \infty$). Our characterization of stochastically stable states is as follows.

Theorem 1. *Consider the logit-response dynamics (with any revision process). A state is stochastically stable if and only if it minimizes $W(s)$ among all states.*

[12] An alternative name for the waste would be *regret*. We prefer to avoid this name for two reasons. First, there is a growing game-theoretic literature where players choose actions according to their associated regret (see e.g. Hart and Mas-Colell (2000)). For us, the waste is rather a technical device and not an objective target. Second, except in the case of asynchronous learning, the waste of a player's choice is only potential regret, since the corresponding payoff will not actually be experienced due to other players simultaneously updating their choices.

The proof, which is relegated to the Appendix, is itself based on an exact characterization of the invariant distribution for finite β (Lemma 2 in the Appendix). Theorem 1 yields a "tree-surgery" technique for the characterization of stochastically stable states of the logit-response dynamics, for arbitrary finite normal-form games and with arbitrary revision processes. It is analogous to the statement that stochastically stable states are those having trees involving a minimal number of mistakes in the mistakes model. In our framework, the number of mistakes is replaced by the sum of payoff losses from transitions which are not possible in the limit as $\beta \to \infty$.

This result makes it possible to focus on minimal waste revision trees to examine stochastic stability. If the set of revising sets is a singleton for every possible transition, as is the case e.g. for asynchronous and instantaneous learning, there is exactly one revision tree per tree involving feasible transitions only, and thus we can directly examine minimal waste trees.

3.2 A Radius-Coradius Result

One of the most powerful results for the actual analysis of models based on the mistakes formulation is the Radius-Coradius theorem due to Ellison (2000). In order to support our tractability claim for the logit model, we now prove a result analogous to Ellison's (2000) Radius-Coradius Theorem in our framework. A directed graph P on S is a *path* if there exists a finite, repetition-free sequence $(s_0, s_1, ..., s_n)$ of states in S with $n = |P|$, such that $(s_m, s_{m+1}) \in P$ and $R_{s_m, s_{m+1}} \neq \emptyset$ for all $m = 0, ..., n-1$. The state s_0 is the *initial* point of the path, the state s_n is the *terminal* point. Since the logit response dynamics is irreducible for any revision process, the set of paths between any two given states is nonempty.

Note that a path as described above is an s_n-tree on the subset of states $\{s_0, ..., s_n\}$ and thus a *revision path* (P, γ) can be simply defined as a particular type of revision tree where P is a path. Denote the set of all revision paths with initial point s and terminal point s' by $\mathcal{P}(s, s')$. The waste $W(P, \gamma)$ of a revision path $(P, \gamma) \in \mathcal{P}(s, s')$ is simply its waste as a revision tree.

The *basin of attraction*[13] of a state s, $B(s) \subseteq S$ is the set of all states s' such that there exists a revision path $(P,\gamma) \in \mathcal{P}(s', s)$ with $W(P,\gamma) = 0$. The *limit set* of state s is the set of states which are connected back-and-forth with s at zero waste, i.e. $L(s) = \{s' \in s | s' \in B(s) \text{ and } s \in B(s')\}$.

The Radius of a state s is defined as

$$R(s) = \min\{W(P,\gamma) | s' \notin B(s), (P,\gamma) \in \mathcal{P}(s, s')\}$$

and is a measure of how easy it is to leave state s. Since the waste is based on payoff differences and not number of mistakes, it takes into account not only the size but also the "depth" of the basin of attraction. The Coradius of s is given by

$$CR(s) = \max_{s' \notin B(s)} \min\{W(P,\gamma) | s'' \in B(s), (P,\gamma) \in \mathcal{P}(s', s'')\}$$

and is a measure of how hard it is to reach s.

Proposition 2. *Suppose a state $s \in S$ is such that $R(s) > CR(s)$. Then, the stochastically stable states are exactly those in $L(s)$.*

Proof. Let $s^* \in S$, $s^* \notin B(s)$. Let $(T^*, \gamma^*) \in \mathcal{T}(s^*)$ solve $\min_{(T,\gamma)\in\mathcal{T}(s^*)} W(T,\gamma)$.

Consider the tree T^* and the complete path from s to s^* in this tree. Since $s^* \notin B(s)$, this path eventually leaves the basin of attraction of s. Let s_1 be the first state in this path which is *not* in $B(s)$. Delete the part of the path from s to s_1. For all states but s that have become disconnected, the fact that they are in $B(s)$ allows to connect them to s (adding the corresponding transitions) with waste zero. If this creates any duplicated edges in the graph, delete the duplicate, but keep only the revising set which ensures waste zero. This saves a waste weakly larger than $R(s)$ (by definition of Radius).

[13] There is a subtle difference between our result and Ellison's (2000). Ellison defines the basin of attraction of a state s as the set of states from which the (unperturbed) dynamics will eventually lead to s with probability one, whereas we define the basin of attraction of s as the set of states such that the unperturbed dynamics (i.e. that involving zero-waste transitions only) leads to s with positive probability.

Add to the revision tree a revision path $(P, \gamma) \in \mathcal{P}(s^*, s)$ which solves $\min\{W(P, \gamma) | s'' \in B(s), (P, \gamma) \in \mathcal{P}(s^*, s'')\}$. Delete any duplicated transitions created when adding (P, γ), keeping the revising sets in γ. This increases the waste by weakly less than $CR(s)$ (by definition of Coradius).

After these two operations we have constructed a new revision tree, rooted in s. If $CR(s) < R(s)$, the total waste has been strictly reduced. It follows that the stochastic potential of s is strictly smaller than the stochastic potential of any s^* not in the basin of attraction of s, thus the latter can not be stochastically stable by Theorem 1.

Consider now a state $s^* \in B(s)$ such that $s \notin B(s^*)$. Consider a minimal-waste s^*-revision tree. Since $s \notin B(s^*)$, in the path connecting s to s^* contained in this tree there exists some transition, say from s_1 to s_2, which causes strictly positive waste. Delete it. Since $s^* \in B(s)$, there exists a zero-waste revision path from s^* to s. Add this path to the revision tree, deleting duplicated transitions. The result is an s_1-revision tree with strictly smaller waste, thus again by Theorem 1, s^* cannot be stochastically stable.

Last, consider any state $s^* \in L(s)$, $s^* \neq s$. Clearly, minimal waste revision trees for both states must have the same waste. In summary, no state out of $L(s)$ can be stochastically stable, but all states in $L(s)$ have the same stochastic potential. Since there are finitely many states, there must exist states with minimum stochastic potential and the conclusion follows. ■

Following Ellison (2000), it is possible to extend this result in two ways. The first would allow to apply the analysis to sets of states rather than a single state. The second would deal with the concept of "modified coradius", which subtracts the radius of intermediate states when computing the coradius, thus providing a more involved but stronger result.

4 Learning in Best-Response Potential Games

In this Section, we illustrate the use of our characterization and provide definite answers to the questions we posed above, that is, to which extent are the findings of convergence to Nash equilibria and selection of potential maximizers robust. To check robustness with respect to the dynamics, we will consider arbitrary revision processes as discussed above. To check robustness with respect to the class of games, we will consider the various generalizations of the concept of potential game.

4.1 Generalized Potential Games

As mentioned in Section 2.2, a finite normal form game $\Gamma = (I, (S_i, u_i)_{i \in I})$ is an (exact) *potential game* if there exists a function $\rho : S \mapsto \mathbb{R}$ (the potential) such that

$$u_i(s_i, s_{-i}) - u_i(s'_i, s_{-i}) = \rho(s_i, s_{-i}) - \rho(s'_i, s_{-i}) \tag{P}$$

for all $i \in I$, $s_i, s'_i \in S_i$, and $s_{-i} \in S_{-i}$. The set of potential maximizers has been shown to be an appealing equilibrium refinement for this class of games. However, it can also be argued that the class of potential games is relatively narrow. Monderer and Shapley (1996) generalized this class as follows. Γ is a *weighted potential game* if (P) is replaced by $u_i(s_i, s_{-i}) - u_i(s'_i, s_{-i}) = w_i\left(\rho(s_i, s_{-i}) - \rho(s'_i, s_{-i})\right)$ for fixed weights $w_i > 0$, $i \in I$. Further, Γ is an *ordinal potential game* if (P) is replaced by the property that $u_i(s_i, s_{-i}) - u_i(s'_i, s_{-i})$ and $\rho(s_i, s_{-i}) - \rho(s'_i, s_{-i})$ have the same sign. Last, Γ is a *generalized ordinal potential game* if (P) is replaced by the property that $u_i(s_i, s_{-i}) - u_i(s'_i, s_{-i}) > 0$ implies that $\rho(s_i, s_{-i}) - \rho(s'_i, s_{-i}) > 0$.

The appeal of generalized ordinal potential games rests on the following characterization. A finite game is generalized ordinal potential if and only if it has the *Finite Improvement Property*, that is, if any path of states generated through unilateral deviations involving strict improvements is necessarily finite.

Obviously, every potential game is a weighted potential game, every weighted potential game is an ordinal potential game, and every ordinal potential game is generalized ordinal potential. Voorneveld (2000) has provided a different generalization of the class of ordinal potential games, and has shown that it is neither included in nor includes the class of generalized ordinal potential games. The game Γ is a *best-response potential game* if there exists a function $\rho^{BR} : S \to \mathbb{R}$ such that $\forall i \in I$, and $s_{-i} \in S_{-i}$,

$$\arg\max_{s_i \in S_i} u_i(s_i, s_{-i}) = \arg\max_{s_i \in S_i} \rho^{BR}(s_i, s_{-i}).$$

Best-response potential games admit a characterization as follows (see Voorneveld (2000, Theorem 3.2)). A normal form game with finitely many players and countable strategy sets is a best-response potential game if and only if any path of states generated through unilateral best responses, and containing at least one strict improvement, is non-cyclic.

4.2 A Convergence Result

We turn now to the question of convergence to Nash equilibria.[14] Theorem 1 allows the following first, immediate observation. In generalized ordinal potential games, convergence to Nash equilibria is not guaranteed, even under asynchronous learning. In other words, non-Nash states can be stochastically stable. To see this, consider the following example.

Example 1. Consider asynchronous learning. The following 2×2 game (left-hand-side table) is Example 4.1.(a) in Voorneveld (2000).

	a	b
a	0,0	0,1
b	0,1	1,0

Payoff Table

	a	b
a	0	1
b	3	2

G.O. Potential

[14]Hofbauer and Sandholm (2002) use stochastic approximation techniques to study convergence of closely-related dynamics to the set of Nash equilibria in potential and supermodular games. The strategy is taking the limit as the population size grows to infinity and the time interval goes to zero, and approximating the paths of the dynamics through a differential equation. In contrast, we study convergence directly on the finite, fixed-population-size, discrete-time dynamics. Baron, Durieu, Haller, and Solal (2002a) have established convergence of the asynchronous logit-response dynamics to *partial* Nash configurations, i.e. strategy profiles where at least one player is choosing a best response.

The only pure-strategy Nash equilibrium is (b, a). This game has a generalized ordinal potential ρ given by the right-hand-side table. However, the game exhibits a best-response cycle, and hence is not a best-response potential game. This best-response cycle contains the links (aa, ab), (ab, bb), (bb, ba), and (ba, aa). Since each of these transitions is a best response for the updating player, we can construct a zero-waste revision tree for all four states. In conclusion, all four states are stochastically stable, even though only one of them is a Nash equilibrium.[15]

This example shows that non-Nash states can be stochastically stable in generalized ordinal potential games under the logit-response dynamics with asynchronous learning. Thus, the next question of interest is *when* does convergence to Nash equilibria obtain. The example above shows that the answer is negative for the class of generalized ordinal potential games.

As an application of Theorem 1, though, we can answer this question in the affirmative for best-response potential games, and hence ordinal potential games. We will also simultaneously perform a robustness check of the convergence result to variations in the way players are chosen to update strategies. Say that a revision process is *regular* if $q_{\{i\}} > 0$ for all $i \in I$. Both standard revision processes in the learning literature, asynchronous learning and independent learning, are clearly regular.

Theorem 2. *If Γ is a finite best-response potential game, the set of stochastically stable states of the logit-response dynamics with any regular revision process is contained in the set of Nash equilibria.*

Proof. Fix a state $s^0 \in S$ which is not a Nash equilibrium of Γ, and hence there exists a coordinate $i \in I$ such that $\max_{s_i \in S_i} u_i(s_i, s^0_{-i}) > u_i(s^0)$. Consider any revision tree $(T^0, \gamma^0) \in \mathcal{T}(s^0)$ with associated waste $W(T^0, \gamma^0)$. Construct a revision tree (T^1, γ^1) from (T^0, γ^0) as follows. Let $s^1 = (s^1_i, s^0_{-i})$ where $s^1_i \in \arg\max_{s_i \in S_i} u_i(s_i, s^0_{-i})$. Add the link (s^0, s^1) with revising set $\{i\}$

[15] In particular, state ab is among the stochastically stable states, even though the set of Nash equilibria is $\{[(p, 1-p), a] \,|0 \leq p \leq \frac{1}{2}\}$, i.e. player 2 never plays strategy b in an equilibrium.

(which is possible by regularity of the revision process) and delete the link (s^1, s^2) leaving s^1 in T^0. The new graph is a tree $T^1 \in \mathcal{T}(s^1)$. The additional transition from s^0 to s^1 causes no waste by definition of s^1. If the contribution of the deleted link (s^1, s^2) to $W(T^0, \gamma^0)$ was positive, $W(T^1, \gamma^1) < W(T^0, \gamma^0)$ holds.[16]

If the contribution was zero and thus $W(T^1, \gamma^1) = W(T^0, \gamma^0)$, proceed as follows. Add a link $(s^1, \hat{s}^2)$ with revising set $\{i\}$ where $\hat{s}^2 = (s_i^2, s_{-i}^1)$ for some $i \in \gamma^0(s^1, s^2)$ such that $s_i^2 \neq s_i^1$. This causes zero waste because (s^1, s^2) caused zero waste. Delete the link $(\hat{s}^2, s^3)$ leaving $\hat{s}^2$ in T^1. The new (labelled) graph is a revision-tree $(T^2, \gamma^2) \in \mathcal{T}(\hat{s}^2)$, with zero waste for the link $(s^1, \hat{s}^2)$.

Iterate the described procedure until deletion of a positive waste link $(\hat{s}^n, s^{n+1})$ occurs, i.e. move along a *best-response compatible* path of states. Since Γ is a best-response potential game, any such path, which started with a strict improvement for a player, is non-cyclic (Voorneveld (2000, Theorem 3.2)), such that iteration actually ends with a revision tree $(T^n, \gamma^n) \in \mathcal{T}(\hat{s}^n)$ where $\hat{s}^n \neq s^0$ and $W(T^n, \gamma^n) < W(T^0, \gamma^0)$. Hence no $(T, \gamma) \in \mathcal{T}(s^0)$ can have minimum waste and, by Theorem 1, s^0 is not stochastically stable. ■

This result generalizes both the class of potential games and the class of logit-response dynamics for which convergence to Nash equilibria obtains. The proof relies crucially on the characterization of finite best-response potential games, that is, the property that any path of states generated through unilateral best responses, containing at least one strict improvement, is non-cyclic. This property is not necessarily fulfilled by generalized ordinal potential games (e.g. it fails in Example 1)

The assumption of regularity of the revision process cannot be dropped. To see this, consider instantaneous learning, where every player receives revision opportunity with probability one. In this case, convergence to Nash equilibria can fail even for exact potential games.

[16] We use the term "waste of a link" as a shortcut for "waste of the revision tree formed by a single link and the chosen revising set".

Example 2. The following 2×2 game is symmetric, and hence an exact potential game (and also a best-response potential game). It has two strict Nash equilibria, (a, a) and (b, b).

	a	b
a	1,1	0,0
b	0,0	1,1

Payoff Table

	a	b
a	1	0
b	0	1

Exact Potential

Under asynchronous learning, both Nash equilibria are stochastically stable, since they both maximize the potential. With our approach it is easy to verify that the same holds under independent learning. Now consider instantaneous learning. Once a Nash state is reached, a waste of 1 is required to leave it, i.e. one of the updating players needs to make a mistake to move to either (a, b) or (b, a).[17] Once the process reaches either (a, b) or (b, a), it alternates between these two states if nobody makes a mistake. Leaving this cycle again causes a waste of 1. Hence the stochastic potential of all states is 2, and they are all stochastically stable. That is, convergence to Nash equilibria might fail even for exact potential games.

4.3 The Irrelevance of Potential Maximizers

This leads us to the second question of interest, namely whether potential maximizers are selected by the logit response dynamics in general. The following example shows that states which globally maximize the potential function of a weighted potential game might fail to be stochastically stable. Thus, although all stochastically stable states of the logit-response dynamics are Nash equilibria for best-response potential games, stochastic stability does not support the use of potential maximizers as an equilibrium refinement for any generalization of potential games, even with asynchronous learning.

[17]Moving directly from one of the Nash states to the other causes a waste of 2, because both players must make a mistake.

Example 3. Let Γ be an asymmetric, pure-coordination, 2×2 game with strategy sets $S_1 = S_2 = \{a, b\}$ and payoffs as given in the following (left-hand-side) table:

	a	b
a	2,2	0,0
b	0,0	10,1

Payoff Table

	a	b
a	2	-6
b	0	4

Weighted Potential

This game has a weighted potential ρ given by the right-hand-side table and weights $w_1 = 1$ and $w_2 = 1/4$. The equilibrium (b, b) is the (unique) potential maximizer.

Consider asynchronous learning. It is straightforward to construct the minimum-waste trees. Note that, since the game is a strict coordination game, states (a, b) and (b, a) can be connected to either of the pure Nash equilibria at zero waste. Thus the minimum waste of a (b, b) tree is equal to the minimum waste necessary to leave (a, a), and vice versa. The waste of the link $(a, a) \mapsto (b, a)$ is $w_1 \cdot (2 - 0) = 2$; the waste of the link $(a, a) \mapsto (a, b)$ is $w_2 \cdot (2 - (-6)) = 2$. Thus the stochastic potential of (b, b) is 2. Consider now state (b, b). The waste of the link $(b, b) \mapsto (a, b)$ is $w_1 \cdot (4 - (-6)) = 10$; the waste of the link $(b, b) \mapsto (b, a)$ is $w_2 \cdot (4 - 0) = 1$. Hence the stochastic potential of (a, a) is 1 and we conclude that (a, a) is stochastically stable, despite not maximizing ρ. This result can also be derived using the Radius-Coradius Theorem. Obviously, $R(a, a) = CR(b, b) = 2$ and $CR(a, a) = R(b, b) = 1$, implying that (a, a) is stochastically stable.

This example shows that the selection of potential maximizers for the asynchronous logit-response is not robust even to slight generalizations of the class of potential games. Now we consider whether the result is robust to generalizations of the class of revision processes.

There are two major differences between the asynchronous-learning case and, say, a process with independent learning. First, the set of revision trees for each state grows, since transitions between any two states become possible. Second, each transition in which not all N players change their action becomes possible via more than one revising set.

Concerning stochastic stability, though, this second issue raises no difficulties. Consider the link (s, s') where the players in J change their action, and assume that a revising set $J' \supset J$ is selected for this transition. It is easy to see that the corresponding waste can only be larger than if the revising set J was selected instead, because sticking to their action might be a non-best response for players in $J' \setminus J$. Hence, when computing the stochastic potential of a state, we can restrict attention to selections for trees that pick the most "parsimonious" revising sets, which prescribes a unique selection for each tree.[18]

The larger set of trees can, however, substantially change other results. We proceed to show that, under independent learning, potential maximizers may fail to be selected even in exact potential games. Thus the result of Corollary 1 is not robust to changes in the specification of revision opportunities either.

Example 4. Consider a $3 \times 3 \times 2$-game with exact potential as given below. Player 1 chooses rows, player 2 chooses columns, and player 3 chooses tables. The payoffs of pure-strategy Nash equilibria are marked by an asterisk.

g

	d	e	f
a	10*	6	0
b	6	0	0
c	0	0	9*

h

	d	e	f
a	0	0	0
b	0	1*	1*
c	0	1*	1

Under asynchronous learning, the potential maximizing state (a, d, g) is stochas-

[18]Essentially, that is the reason why Theorem 2 holds for any regular revision process.

tically stable by Corollary 1. Consider independent learning instead. The basin of attraction of state (c, f, g) contains all states except (a, d, g), (a, e, g) and (b, d, g). Any minimal waste path from (c, f, g) to one of these states, for example the path $((c, f, g), (a, f, g), (a, d, g))$ or $((c, f, g), (c, f, h), (c, d, h), (b, d, g))$ is associated with a waste of 9, such that $R(c, f, g) = 9$. The transition $((a, d, g), (b, e, g))$, though, has an associated waste of only 8 when players 1 and 2 switch simultaneously. The states (a, e, g) and (b, d, g) can be connected to $B(c, f, g)$ at an even lower waste, such that $CR(c, f, g) = 8$. Proposition 2 then implies that (c, f, g) is stochastically stable, despite the fact that it does not maximize the exact potential.

5 Generalizations and Extensions

Although we have focused on the logit-response dynamics, our approach to stochastic stability is susceptible of generalization to a wider class of learning processes. In this Section, we briefly report on this generalization.

Consider a Markov chain $\{X_t\}_{t\in\mathbb{N}}$ on a finite state space Ω. Denote the stationary transition probabilities by $P_{\omega,\omega'} = Pr(X_t = \omega'|X_{t-1} = \omega)$. A *transition mechanism* from state ω is a mapping $Q : \Omega \to \mathbb{R}_+$ such that $Q(\omega') > 0$ for at least some $\omega' \in \Omega$. The interpretation is that from a given state ω, there might be different, alternative processes giving rise to a transition to other states. Conditional on the transition mechanism Q being selected, a state $\omega' \in \Omega$ will be reached from ω with probability

$$Q(\omega')/\sum_{\omega''\in\Omega} Q(\omega'').$$

Denote by M_ω the set of transition mechanisms available at ω, and let $M = \bigcup_{\omega\in\Omega} M_\omega$. Note that the sets M_ω need not be pairwise disjoint, so that a transition mechanism might be available at several or even all states (e.g. a random mutation). Further, let $M_{\omega,\omega'} = \{Q \in M_\omega|Q(\omega') > 0\}$, i.e. the set of mechanisms which are available at ω and may lead to ω'.

Definition 4. Let X_t be a Markov chain on the finite state space Ω. A *decomposition* of X_t is a tuple $(M_\omega, q_\omega)_{\omega\in\Omega}$ such that, for each $\omega \in \Omega$,

(i) M_ω is a nonempty, finite set of transition mechanisms,

(ii) $q_\omega \in \Delta M_\omega$ is a full-support probability measure on M_ω, and

(iii) for each $\omega' \in \Omega$,

$$P_{\omega,\omega'} = \sum_{Q\in M_{\omega,\omega'}} q_\omega(Q) \frac{Q(\omega')}{\sum_{\omega''\in\Omega} Q(\omega'')}.$$

Obviously, any finite Markov chain admits a trivial (and not very useful) decomposition with $M_\omega = \{Q_\omega\}$ and $Q_\omega(\omega') = P_{\omega,\omega'}$ for all ω'.

Definition 5. A *log-linear Markov family* is a family of finite Markov chains X_t^β with $\beta \in [1, +\infty[$, defined on a common state space Ω, such that

(i) the chain $X_t = X_t^1$ is irreducible and admits a decomposition $(M_\omega, q_\omega)_{\omega\in\Omega}$,

(ii) each X_t^β with $\beta > 1$ admits a decomposition $(M_\omega^\beta, q_\omega)_{\omega\in\Omega}$ given by

$$M_\omega^\beta = \{Q_\omega^\beta | Q_\omega \in M_\omega\}$$

where $\ln Q_\omega^\beta(\omega') = \beta \cdot \ln Q_\omega(\omega')$ whenever $Q_\omega(\omega') > 0$ (and $Q_\omega^\beta(\omega') = 0$ otherwise).

A log-linear Markov family can be seen as an interpolation between the X_t^1 chain (the "pure noise" chain) and a "limit chain" as $\beta \to \infty$. Irreducibility of the pure-noise chain implies irreducibility of all chains in the family, but not of the limit chain. A state ω is *stochastically stable* if $\lim_{\beta\to\infty} \mu^\beta(\omega) > 0$, where μ^β is the invariant distribution for $\beta > 0$.

Example 5. Consider the logit-response dynamics with revision process q. Its decomposition corresponds to equation 4. That is, the transition mechanisms Q_J available at a state s correspond to the revising sets J, and $Q_J(s') = e^{\beta \cdot U_J(s',s)}$. The pure-noise chain corresponds to the $\beta = 1$ case, and the limit chain is the best-response dynamics.

Given a log-linear family, a *transition tree* is defined analogously to a revision tree, i.e. a pair (T, γ) where T is a tree such that $(\omega, \omega') \in T$ only if $M_{\omega,\omega'} \neq \emptyset$ and $\gamma : T \mapsto M$ is such that $\gamma(\omega, \omega') \in M_{\omega,\omega'}$ for each $(\omega, \omega') \in T$. That is, γ selects a transition mechanism for each link in the tree. Denote the set of all transition ω-trees by $\mathcal{T}(\omega)$.

Analogously to Lemma 2 in the Appendix, straightforward but cumbersome computations allow to give an exact characterization of the invariant distribution $\mu^\beta(\omega)$. This in turn allows to establish the analogue of Theorem 2. The *waste* of a revision tree $(T, \gamma) \in \mathcal{T}(\omega)$ is defined as

$$W(T, \gamma) = \sum_{(\omega,\omega') \in T} \left(\max_{\omega'' \in \Omega} Q_{\gamma(\omega,\omega')}(\omega'', \omega) \right) - Q_{\gamma(\omega,\omega')}(\omega', \omega).$$

The *stochastic potential* of a state ω is defined as $W(\omega) = \min_{(T,\gamma) \in \mathcal{T}(\omega)} W(T, \gamma)$.

Theorem 3. *Consider a log-linear Markov family. A state ω is stochastically stable if and only if it minimizes $W(\omega)$ among states.*

Log-linear Markov families can be used to analyze a large variety of learning models. In the case of the logit-response dynamics, transition mechanisms correspond to different groups of players who are updating at the same time. Transition mechanisms can, however, also be used to model alternative behavioral rules of the agents, such as imitation. Varying memory length, possibly correlated with the complexity of observed histories, or differences in observability of the others' actions across players, states and points in time, could all be captured through appropriately defined transition mechanisms.

In this paper, we have focused on the logit-response dynamics and hence it is natural to consider log-linear Markov families as a generalization. It would of course be possible to further generalize the framework to allow for perturbations which are not of the log-linear form. Such a framework would allow to encompass e.g. the mistakes model as a particular case (with the pure noise chain being the mutation process and the limit chain myopic best reply). Related approaches have been pursued by Myatt and Wallace (2003) and Beggs (2005), who consider families of Markov chains with transition probabilities P^β such that the limits $\lim_{\beta\to\infty} -\frac{1}{\beta}\ln P^\beta_{\omega,\omega'}$ are well-defined.[19]

6 Conclusions

The mistakes model of Kandori, Mailath, and Rob (1993) and Young (1993) is analytically flexible due to the well-known graph-theoretic characterization of the stochastically stable states. It has been often criticized, e.g. by Bergin and Lipman (1996), due to the sensitivity of the results to the specification of the noise process. Other dynamics, like the logit-response dynamics of Blume (1993), present more solid foundations but analytical results can be derived only for particularly convenient frameworks.

Here we have presented a characterization of the stochastically stable states of a generalization of the logit-response dynamics. This new characterization is in the spirit of the mistakes model. We have illustrated the approach studying convergence to the set of Nash equilibria of the logit-response dynamics in general classes of games. Convergence obtains for best-response potential games but fails for generalized ordinal potential games. The selection of potential maximizers in exact potential games appears to be a fragile result, robust neither to generalizations of the considered game class nor to the specification of revision opportunities.

[19]Myatt and Wallace (2003) examine stochastic stability in a learning model where payoffs are perturbed by normally distributed shocks. They show that the addition of a strictly dominated strategy can change the selection result. Following the approach in Ellison (2000), Beggs (2005) uses graph-theoretic arguments to obtain general results on waiting times. See also Dokumaci and Sandholm (2007). Hofbauer and Sandholm (2007) examine perturbed best-response dynamics (including logit), but concentrate on the large-population limit.

A Proof of Theorem 1

The proof proceeds as follows. First, we introduce a few auxiliary concepts. Then we use these concepts to provide an exact (but cumbersome) characterization of the invariant distribution for fixed, finite β in Lemma 2. Last, we use this characterization to prove Theorem 1.

Given a graph G on the state space S, a mapping

$$\gamma : G \mapsto \mathcal{P}(I)$$

such that $\gamma(s, s') \in R_{s,s'}$ for each $(s, s') \in G$ is called a *revision selection for G*. For each transition in G, a revision selection for G picks exactly one of the possible revising sets making that transition (potentially) possible. Thus, a revision tree is a pair (T, γ) made of an s-tree involving only feasible transitions under the revision process q, and a revision selection γ for T.

Denote the set of all revision selections for a graph G by $\mathcal{S}(G)$.

Let $M = S \times \mathcal{R}^q$ denote the set of all pairs made of one state and one revising set. Consider a subset $N \subseteq M$. A *realization* r for N is a mapping $r : N \mapsto S$ such that $r(s, J) \in R_s^J$ for all $s \in S$ and all $J \in \mathcal{R}^q$, $J \neq \emptyset$. The set of all realizations for N is denoted $R(N)$. A *complete realization* is just a realization for M.

A *completion* of a revision tree (T, γ) is a complete realization such that $r(s, \gamma(s, s')) = s'$ for all $(s, s') \in T$. In words, a completion assigns a feasible outcome for each state and each possible revising set such that, whenever the revising set is the one specified by the selection for the (unique) arrow leaving the state in the tree, the outcome is precisely the state this arrow leads to. Let $\mathcal{C}(T, \gamma)$ be the set of all completions of (T, γ).

If γ is a revision selection for a tree T and $N^\gamma = \{(s, \gamma(s, s')) | (s, s') \in T\}$, then $R(N^\gamma)$ can be interpreted as the set of possible realizations of the selection γ.

Lemma 2 (The Decomposition Lemma)**.** *The invariant distribution* μ *of the logit-response dynamics with revision process* q *satisfies for each* $s^* \in S$

$$\mu(s^*) \propto \sum_{(T,\gamma)\in\mathcal{T}(s^*)} P(T,\gamma) \sum_{r\in\mathcal{C}(T,\gamma)} e^{Q(r)}, \tag{5}$$

where

$$P(T,\gamma) = \prod_{(s,s')\in T} q_{\gamma(s,s')}$$

and

$$Q(r) = \sum_{s\in S}\sum_{J\in\mathcal{R}^q} U_J(r(s,J),s).$$

Proof. Let $\mathcal{T}_0(s^*)$ be the set of all s^*-trees. By Freidlin and Wentzell (1988, Lemma 3.1),

$$\mu(s^*) \propto \sum_{T\in\mathcal{T}_0(s^*)} \prod_{(s,s')\in T} P_{s,s'}.$$

Note that s-trees including transitions which are not feasible under q contribute zero to the sum above. Fix a tree $T \in \mathcal{T}_0(s^*)$ such that all transitions are feasible under q. Using the decomposition of the transition probabilities (4),

$$\prod_{(s,s')\in T} P_{s,s'} = \prod_{(s,s')\in T}\left(\sum_{J\in R_{s,s'}} q_J \frac{e^{U_J(s',s)}}{\sum_{s''\in R_s^J} e^{U_J(s'',s)}}\right).$$

Expanding the RHS yields

$$\sum_{\gamma\in\mathcal{S}(T)}\left(\prod_{(s,s')\in T} q_{\gamma(s,s')} \frac{e^{U_{\gamma(s,s')}(s',s)}}{\sum_{s''\in R_s^{\gamma(s,s')}} e^{U_{\gamma(s,s')}(s'',s)}}\right) =$$

$$= \sum_{\gamma\in\mathcal{S}(T)}\left(P(T,\gamma)\frac{e^{\sum_{(s,s')\in T} U_{\gamma(s,s')}(s',s)}}{\prod_{(s,s')\in T}\sum_{s''\in R_s^{\gamma(s,s')}} e^{U_{\gamma(s,s')}(s'',s)}}\right).$$

Expanding the denominator in the last expression yields

$$\sum_{\gamma\in\mathcal{S}(T)}\left[P(T,\gamma)\, e^{\sum_{(s,s')\in T} U_{\gamma(s,s')}(s',s)}\left(\sum_{r\in R(N^\gamma)} e^{\sum_{(s,s')\in T} U_{\gamma(s,s')}(r(s,\gamma(s,s')),s)}\right)^{-1}\right].$$

We multiply and divide the last expression by

$$\sum_{r\in R(M\setminus N^{\gamma})} e^{\sum_{(s,J)\in M\setminus N^{\gamma}} U_J(r(s,J),s)}$$

and obtain

$$\sum_{\gamma\in\mathcal{S}(T)} \left[P(T,\gamma) \left(\sum_{r\in\mathcal{C}(T,\gamma)} e^{\sum_{(s,J)\in M} U_J(r(s,J),s)} \right) \left(\sum_{r\in R(M)} e^{\sum_{(s,J)\in M} U_J(r(s,J),s)} \right)^{-1} \right].$$

The last term in brackets is a constant which is independent of both γ and T, and hence is irrelevant for proportionality of $\mu(s^*)$. The proof is completed observing that $\sum_{(s,J)\in M} U_J(r(s,J),s) = Q(r)$. ■

We are now ready to prove Theorem 1, i.e. stochastically stable states are those where the waste is minimized across revision trees.

Proof of Theorem 1. Fix a revision process q. Let μ^{β} denote the invariant distribution of the logit response dynamics for noise level β. By Lemma 2 we have that, for every state s,

$$\mu^{\beta}(s) \propto \sum_{(T,\gamma)\in\mathcal{T}(s)} P(T,\gamma) \sum_{r\in\mathcal{C}(T,\gamma)} e^{\beta\cdot Q(r)},$$

where $Q(r) = \sum_{s\in S}\sum_{J\in\mathcal{R}^q} U_J(r(s,J),s)$, and $P(T,\gamma) > 0$ for all (T,γ). As $\beta\to\infty$, only the completion r which maximizes $Q(r)$ among all completions for all revision trees $(T,\gamma)\in\mathcal{T}(s)$ matters for stochastic stability of state s, since its effect dominates for large β. Specifically, stochastically stable states, i.e. those satisfying that $\lim_{\beta\to\infty}\mu^{\beta}(s) > 0$, are exactly those states $s\in S$ for which the expression

$$\max_{(T,\gamma)\in\mathcal{T}(s)} \max_{r\in\mathcal{C}(T,\gamma)} Q(r)$$

is maximal among all states.

For any given revision tree (T,γ), the completion r^* which maximizes $Q(r)$ among all completions $r\in\mathcal{C}(T,\gamma)$ clearly involves only best responses for all revising players on all pairs $(s, J) \notin N^\gamma$, i.e. in state-revising set pairs not used for transitions in the revision tree.

Let $r^{\max}$ be a complete realization involving only best responses. It follows that a state $s \in S$ maximizes $\max_{(T,\gamma)\in\mathcal{T}(s)} \max_{r\in\mathcal{C}(T,\gamma)} Q(r)$ if and only if it maximizes $\max_{(T,\gamma)\in\mathcal{T}(s)} \sum_{(s,s')\in T} \big(U_{\gamma(s,s')}(s', s) - U_{\gamma(s,s')}(r^{\max}(s, \gamma(s, s')), s)\big)$. Since $\sum_{(s,s')\in T} \big(U_{\gamma(s,s')}(s', s) - U_{\gamma(s,s')}(r^{\max}(s, \gamma(s, s')), s)\big) = -W(T, \gamma)$, it follows that stochastically stable states are those having minimal stochastic potential $\min_{(T,\gamma)\in\mathcal{T}(s)} W(T, \gamma)$. ■

References

Alós-Ferrer, C., and A. B. Ania (2005): "The Evolutionary Stability of Perfectly Competitive Behavior," *Economic Theory*, 26, 497–516.

Alós-Ferrer, C., and G. Kirchsteiger (2007): "Learning and Market Clearing," Mimeo.

Baron, R., J. Durieu, H. Haller, and P. Solal (2002a): "A Note on Control Costs and Logit Rules for Strategic Games," *Journal of Evolutionary Economics*, 12, 563–575.

——— (2002b): "Control Costs and Potential Functions for Spatial Games," *International Journal of Game Theory*, 31, 541–561.

Beggs, A. (2005): "Waiting Times and Equilibrium Selection," *Economic Theory*, 25, 599–628.

Bergin, J., and B. L. Lipman (1996): "Evolution with State-Dependent Mutations," *Econometrica*, 64, 943–956.

Blume, L. (1993): "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5, 387–424.

Blume, L. (1997): "Population Games," in *The Economy as an Evolving Complex System II*, ed. by B. Arthur, S. Durlauf, and D. Lane, pp. 425–460. Addison-Wesley, Reading, MA.

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

Dokumaci, E., and W. H. Sandholm (2007): "Stochastic Evolution with Perturbed Payoffs and Rapid Play," Mimeo, University of Wisconsin-Madison.

Ellison, G. (2000): "Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution," *Review of Economic Studies*, 67, 17–45.

Freidlin, M., and A. Wentzell (1988): *Random Perturbations of Dynamical Systems, 2nd Ed.* Springer Verlag, New York.

Hart, S., and A. Mas-Colell (2000): "A Simple Adaptive Procedure Leading to Correlated Equilibrium," *Econometrica*, 68, 1127–1150.

Hofbauer, J., and W. Sandholm (2002): "On the Global Convergence of Stochastic Fictitious Play," *Econometrica*, 70, 2265–2294.

——— (2007): "Evolution in Games with Randomly Disturbed Payoffs," *Journal of Economic Theory*, 132, 47–69.

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

Hofbauer, J., and G. Sorger (1999): "Perfect Foresight and Equilibrium Selection in Symmetric Potential Games," *Journal of Economic Theory*, 85, 1–23.

Kandori, M., G. J. Mailath, and R. Rob (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 61, 29–56.

Kandori, M., and R. Rob (1995): "Evolution of Equilibria in the Long Run: A General Theory and Applications," *Journal of Economic Theory*, 65, 383–414.

Maruta, T. (2002): "Binary Games with State Dependent Stochastic Choice," *Journal of Economic Theory*, 103(2), 351–376.

Mattsson, L.-G., and J. W. Weibull (2002): "Probabilistic Choice and Procedurally Bounded Rationality," *Games and Economic Behavior*, 41(1), 61–78.

McKelvey, R. D., and T. R. Palfrey (1995): "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10(1), 6–38.

Monderer, D., and L. Shapley (1996): "Potential Games," *Games and Economic Behavior*, 14, 124–143.

Myatt, D. P., and C. Wallace (2003): "A Multinomial Probit Model of Stochastic Evolution," *Journal of Economic Theory*, 113, 286–301.

Sandholm, W. H. (1998): "Simple and Clever Decision Rules for a Model of Evolution," *Economics Letters*, 61, 165–170.

——— (2007): "Pigouvian Pricing and Stochastic Evolutionary Implementation," *Journal of Economic Theory*, 132, 367–382.

Thurstone, L. (1927): "A Law of Comparative Judgement," *Psychological Review*, 34, 273–286.

van Damme, E., and J. W. Weibull (2002): "Evolution in Games with Endogenous Mistake Probabilities," *Journal of Economic Theory*, 106, 296–315.

Vega-Redondo, F. (1997): "The Evolution of Walrasian Behavior," *Econometrica*, 65, 375–384.

Voorneveld, M. (2000): "Best-Response Potential Games," *Economics Letters*, 66, 289–295.

Young, P. (1993): "The Evolution of Conventions," *Econometrica*, 61, 57–84.

——— (1998): *Individual Strategy and Social Structure*. Princeton University Press, Princeton, New Jersey.

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Hauptstr. 90
CH-8280 Kreuzlingen 2

Telefon: +41 (0)71 677 05 10
Telefax: +41 (0)71 677 05 11

info@twi-kreuzlingen.ch
www.twi-kreuzlingen.ch