On the Evolution of Beliefs

Samuel Häfner*

May 29, 2014

Abstract

This paper explores the evolutionary foundations of belief formation in strategic interactions. The framework is one of best response dynamics in normal form games where the revising agents form stochastic beliefs about the actual strategy distribution in the population. The shares of agents drawing from the same belief distribution are subject to the replicator dynamics. The basic idea is that beliefs translate into behavior, behavior translates into fitness, and fitness then determines the evolutionary success of a belief distribution. A belief distribution is called replicator dynamics stable if – given that all agents in the population draw their beliefs from that distribution – any small share of an intruding belief distribution is crowded out again. We show how this notion relates to the traditional replicator stability of strategies, and how the framework can be applied to study the evolutionary stability of sampling procedures, and the stability of mixed equilibria in asymmetric normal form games.

Keywords: Best Response Dynamics, Sampling, Evolution, Normal Form Games, Two-Speed Dynamics
JEL-Classifications: C73, D83

*University of Basel, Faculty of Business and Economics, Peter Merian-Weg 6, CH-4002 Basel, Switzerland, samuel.haefner@unibas.ch.

1 Introduction

The rational choice paradigm explains strategic behavior by means of preferences, beliefs, and actions: Agents consistently choose from the actions available in order to maximize their utility given the belief that they hold about the action choice of others. Economic theorists engaged in analyzing the evolutionary foundations of strategic behavior have so far mainly focused on preferences, and actions. However, there is no reason to think that the third component of the paradigm, that is, beliefs that agents hold, or the way agents come to hold their beliefs should not be subject to evolutionary forces, either.

This paper sketches a route to analyze evolutionary stable beliefs in strategic interactions, and to characterize the behavior that these beliefs induce. The basic idea is similar to the idea in the indirect evolutionary approach that is usually applied to preferences (cf. Güth & Yaari 1992): Beliefs determine behavior, behavior determines fitness, and fitness in turn determines the evolutionary success of the beliefs. Other than in the classical approach that employs a static evolutionary concept, however, we analyze this causal chain by spelling out an explicit model of both the dynamics describing how agents arrive at equilibrium play given their beliefs, and of the evolutionary dynamics describing how the beliefs in the population evolve.

We assume a best response dynamics framework: Individual agents of a unit mass population are repeatedly and randomly matched in a 2×2 normal form game. Agents myopically play best response to their belief about the current strategy distribution in the population, but from time to time, a subset of the agents gets the opportunity to revise their belief. Other than in the standard best response dynamics where the revised belief corresponds to the actual strategy distribution at the time of revision (cf. Gilboa & Matsui 1991), we here assume that revising agents draw a stochastic belief about the strategy distribution from a belief distribution that is conditional on the actual strategy distribution. The conception of random beliefs is motivated by the observation that perception – being the device that generates an agents beliefs about his environment – is a partly random process (cf. Kahneman & Tversky 1996).

Randomness in perception brings about an indeterminacy to the best response dynamics in the sense that any strategy distribution can be chosen to be a steady state of the best response dynamics with appropriately specified conditional belief distributions. This observation motivates our model's second building block. We resort to evolution as a selection device between belief distributions. We present a model in which differing belief distributions from a given family of possible belief distributions compete within a population: We divide the population into two sub-populations, let agents of a given sub-population draw from the same belief distribution, and put the standard replicator dynamics on the sub-population shares. Under this dynamics, the sub-population share of those agents drawing their beliefs from a distribution that yields a fitness level that is higher than the population average fitness level grows, whereas sub-population share of those agents drawing their beliefs from a distribution that yields a fitness level that is lower than the population average fitness level shrinks. We focus on states with belief-monomorphic populations – i.e. states where all agents in a population draw from the same belief distributions – and analyze the behavior of the system after an intrusions by a tiny share of agents drawing from a different belief distribution.

The key to our analysis lies in understanding the dynamical system comprising the best response dynamics and the replicator dynamics as a slow-fast system. Slow-fast systems are characterized by two different time-scales. The variables subject to the faster time-scale are taken together in what is called the fast node, and the others in the slow node, respectively (c.f. Berglund & Gentz 2006 for an introduction). In the context of this paper, we will understand the best response process as the fast, and evolution as the slow node.

Two-speed dynamics have been employed in models of preference evolution, albeit only implicitly: For example, Sandholm (2001*b*), Dekel, Ely & Yilankaya (2007), or Alger & Weibull (2013) deal with the limit case of play adapting infinitely fast to changes in the distribution of preferences. They do so by assuming that agents are aware of the changing nature of the preferences, and always play equilibrium given the current distribution of preferences.[1] In this paper, we resort to a well-known result in the context of slow-fast systems known as Tykhonov's Theorem (cf. Kokotovic 1984, Theorem 2.1), that states sufficient conditions to approximate the behavior of the slow-fast system by the solution of a reduced system, and thus allows the analysis of the less extreme cases where the adaptation of play is explicitly modeled and, crucially, does not happen infinitely fast.

Motivated by Tykhonov's Theorem, we say that a belief distribution is stable under the replicator dynamics at a strategy distribution, if in a state with all agents of the population drawing their beliefs from the same distribution and the best response dynamics at rest yielding that strategy distribution, there is no other belief distribution yielding a differing distribution of actions that can intrude beyond possibly a tiny share whose size vanishes in the speed of the best response dynamics. Tykhonov's Theorem yields conditions both on the best response dynamics as well as a on the replicator dynamics such that a belief distribution is replicator dynamics stable. These separate conditions are

[1] Sandholm (2001*b*) discusses relaxing the assumption of infinitely fast adapting play in his setup. See the literature section for a more thorough discussion of his approach and how it relates to ours.

very convenient as the full slow-fast system is analytically intractable. We also briefly discuss the close connection of our notion of stability to the classical notion of replicator dynamics stability of strategies (cf. Weibull 1997).

We present two applications of our approach: First, we look at beliefs that are generated by random strategy sampling in the population. By this, we mean that when agents revise their belief about the current population state, they do so by randomly sampling agents from the population and therefrom deducing a point estimate by looking at the strategy distribution in the sample. We can show that in games of the Hawk-Dove variety, sampling procedures that lead to replicator dynamics stable belief distributions are generically biased. In a second application, we extend the framework to asymmetric 2×2-normal form games, and look at games of the Matching Pennies variety. Here, we can show that the unique population state brought about by replicator dynamics stable belief distributions corresponds to the (unique) mixed strategy Nash equilibrium distribution.

The paper is organized as follows: Section 2 presents the basic model for symmetric 2×2-normal form games, and sets up the dynamics to be analyzed. Section 3 elaborates on the stability criterion, states the main result giving conditions for stability, and relates our notion of stability to the classical notion of replicator dynamics stability of strategies. In Section 4, we look at a the special class of belief distributions generated by random sampling. In Section 5, the main theorem is applied to a generalized 2×2 matching pennies game. Section 6 discusses the relation to the literature, and the last section concludes.

2 The Basic Model

We consider the following setup: There is one population whose agents are repeatedly randomly pairwise matched in a symmetric 2×2-normal form game. We label the two strategies by strategy 1, and 2, respectively. We write $z \in [0,1]$ for the share of agents playing strategy 1, and call it the population state. The expected utility in a match to an agent playing strategy $i \in 1,2$ is given by $U_i(z)$. As the dynamics to be set up only depend on payoff difference between strategies, and these differences are invariant to shifts in the utilities for a given opponent strategy, we can normalize these payoffs as follows (cf. p.40 in Weibull 1997): $U_1(z) = az$, and $U_2(z) = b(1-z)$, where $a, b \in \mathbb{R}$.

We restrict attention to cases of $\text{sgn}(a) = \text{sgn}(b)$. With $a, b > 0$, we have a game of the Coordination Game variety; with $a, b < 0$, we have a game of the Hawk-Dove variety. The reason for the sign-restriction is the following: With $a > 0 > b$, or with $a < 0 < b$ (corresponding to games of the Prisoner's Dilemma variety) we have a strictly dominant strategy, and hence the belief about the other players' strategy choices (which is the main interest in this paper) is irrelevant.

Agents hold individual beliefs $\hat{z} \in [0,1]$ about the current population state z. We assume that agents take their environment for stationary, that is, they neither update their belief based on the history of play, nor do they take into account what influence their choice of action might have on the action choice others. In every match, the agents play a best response to their beliefs. We assume that an agent plays strategy 1 whenever this yields a weakly higher utility than strategy 2.[2] Accordingly, we call $\mathscr{B} \equiv \{z : U_1(z) \geq u_2(z)\}$ the best response set of the agents. The set $\mathscr{B}$ contains those population states z to which agents have strategy 1 as a best response. Agents with a belief $\hat{z} \in \mathscr{B}$ play strategy 1, whereas agents with a belief $\hat{z} \in [0,1] \setminus \mathscr{B}$ play strategy 2. That is, if $a, b < 0$, we have $\mathscr{B} = [0, b/(a+b)]$, and if $a, b > 0$, we have $\mathscr{B} = [b/(a+b), 1]$.

From time to time, agents get the opportunity to revise their belief about the current population state. For each agent, these opportunities occur according to a Poisson process, and arrive with rate $\lambda > 0$. That is, as we look at ever shorter time intervals of length $\mathrm{d}\tau > 0$, the probability that an agent receives an opportunity to revise his beliefs in the period $[t, t + \mathrm{d}\tau)$ is given by $\lambda \mathrm{d}\tau$. A revising agent forms a new belief $\hat{z}$ about the current population state, and plays best response to that belief until the next opportunity to revise arises.

The revision stage is modeled in the following reduced way: An agent revising at a current population state $z \in [0,1]$ forms a revised belief $\hat{z}$ that is a draw from a conditional cumulative distribution function $F(.|z) : [0,1] \rightarrow [0,1]$. If $F(.|z)$ has a density, we denote it by $f(.|z)$. If $F(.|z)$ is discrete, we accordingly interpret $f(.|z)$ as a probability. The belief distribution F belongs to the family $\mathscr{F}$ of possible belief distributions which is a subset of the set of all distribution functions on $[0,1]$ that are conditional on $z \in [0,1]$. While the specific characterization of $\mathscr{F}$ will depend on the respective application considered, we make two assumptions about $\mathscr{F}$ that hold throughout the paper:

Assumption 1 (Differentiability)**.** *It holds for all members F of $\mathscr{F}$ that, $\forall x, z \in [0,1]$, $F(x|z)$ is continuously differentiable in z.*

Throughout the paper, we say that a function is differentiable on a closed convex subset of the Euclidean space if the respective one-sided derivatives exist at all boundary points of the set. The assumption guarantees that the beliefs of agents do not change too abruptly when the actual population state changes. On a technical level, differentiability of F in z is required in order that the vector field of the dynamics to be set up is differentiable.

[2] The tie-breaking rule in case of indifference is not important. All the results go through for any (probabilistic) tie-breaking rule.

Assumption 2 (Richness). *For every $u, v, z \in [0,1]$, $\exists F \in \mathscr{F}$ such that $u = F(v|z)$.*

This assumption states that at any population state z and for any arbitrary $u, v \in [0,1]$, we always have at least one distribution $F \in \mathscr{F}$ under which the probability that the belief is below v is given by u. This will ensure that any population state z can be supported as a steady-state of the dynamics to be set up.

In the general description of the model, we are not interested in the way the belief distributions in $\mathscr{F}$ come about. Nevertheless, it is perhaps natural to think that there are underlying information gathering and processing procedures that yield specific characterizations of the members of $\mathscr{F}$. One example of such a procedure, that we refer to repeatedly in the following and discuss in some more detail in Section 4, is random sampling:

Example 1 (Random Sampling). Under random sampling, the revising agent draws a finite random sample of players from the population, observes the strategies played by the agents in the sample, and then forms a point estimate of the current state z. Let the sample size be $M \in \mathbb{N}_{++}$, and let $m \leq M$ be the number of strategy 1 agents observed in the sample. The belief about z is then given by $\hat{z} = m/M$. The belief $\hat{z}$ is binomially distributed, and a typical member of the family $\mathscr{F}$ induced by such sampling procedures is given by

$$F(x|z) = \sum_{i=1}^{\lfloor M \cdot x \rfloor} \binom{M}{i} z^i (1-z)^{M-i}$$

with $x, z \in [0,1]$, and with $\lfloor x \rfloor$ denoting the floor operator that returns the highest integer weakly below x. Note that a family $\mathscr{F}$ that merely contains belief distributions F that are generated by random sampling with finite sample sizes is not rich. In Section 4, we suitably extend the sampling procedure to what we call generalized sampling, which will guarantee that the resulting family $\mathscr{F}$ of belief distributions is indeed rich.

In order to set up the dynamics, we divide the population in two sub-populations labeled 1 and 2. The sub-populations differ in the belief distributions that revising agents draw new beliefs from. We denote the belief distribution of sub-population $q = 1, 2$ by $F_q \in \mathscr{F}$. The mass of sub-population 1 is denoted by $\varphi \in [0,1]$. We call the population distribution-monomorphic if $\varphi \in \{0,1\}$: The value $\varphi = 0$ corresponds to the case where all agents in the population draw from belief distribution F_2, while $\varphi = 1$ corresponds to the case where all agents draw from belief distribution F_1. Strategy shares within sub-populations are denoted by $\mu_q \in [0,1]$, such that the current population state z is given by $z = \varphi\mu_1 + (1-\varphi)\mu_2$. We sometimes write z as a function $z(\mu, \varphi)$, $\mu = (\mu_1, \mu_2)$, in order to emphasize its dependence on (μ, φ).

2.1 Best Response Dynamics

We first look at the best response dynamics of the population state z for a fixed sub-population share $\varphi \in [0,1]$. To set up the dynamics, we start with discrete time, and then take the continuous time limit. Let

$$\zeta_q(z) \equiv \int_{\hat{z}\in\mathscr{B}} dF_q(\hat{z}|z)$$

be the reaction function of an agent of sub-population $q = 1,2$ that specifies the probability that an agent chooses strategy 1 after revision at a population state $z \in [0,1]$ when drawing a belief from a distribution F_q. Note that $\zeta_q(z)$ is differentiable in z because $F_q(x|z)$ is differentiable in z by Assumption 1.

We twice resort to a standard abuse of the law of large numbers, and assume (i) that the fraction of agents revising their belief in a short time period of length $\mathrm{d}\tau > 0$ is given by $\delta\mathrm{d}\tau$, and (ii) that the fraction of revising agents in sub-population q choosing strategy 1 at population state z can be written as $\zeta_q(z)$. This then allows us to write the sub-population strategy shares in sub-population q at time $t+\mathrm{d}\tau$ as $\mu_q(t+\mathrm{d}\tau) = (1-\lambda\mathrm{d}\tau)\mu_q(t) + \lambda\mathrm{d}\tau\zeta_q(z(t))$. Taking the continuous time limit, that is, letting $\mathrm{d}\tau \to 0$, the dynamics of the sub-population strategy shares can be written as:

$$\frac{\mathrm{d}\mu_q(t)}{\mathrm{d}t} = \lambda\left(\zeta_q(z(t)) - \mu_q(t)\right), \quad q = 1,2 \tag{1}$$

Letting $\zeta = (\zeta_1, \zeta_2)$ be the vector of the reaction functions of the sub-populations, we re-express the best response process compactly as

$$\frac{\mathrm{d}\mu(t)}{\mathrm{d}t} = \lambda\left[\zeta\left(z(\mu(t),\varphi(t))\right) - \mu(t)\right] \equiv g\left(\mu(t),\varphi(t)\right), \quad \mu \in [0,1]^2 \tag{2}$$

We refer to system (2) as the best response node. The following lemma establishes existence of at least one fixed point of the best response node for any two belief distributions $F_1, F_2 \in \mathscr{F}$ and any share φ of belief distributions in the populations:

Lemma 1. *Fix $F_1, F_2 \in \mathscr{F}$. Then, $\forall\varphi \in [0,1]$, the set of roots $\mu \in [0,1]^2$ of $g(\mu,\varphi)$,*

$$\Upsilon^*(\varphi) \equiv \left\{\mu \in [0,1]^2 : \mu = \zeta\left(z(\mu,\varphi)\right)\right\}, \tag{3}$$

is non-empty.

Proof. By definition, $z(\mu,\varphi)$ is linear in the elements of μ, and hence continuous in μ. From Assumption 1 it follows that ζ is differentiable, and hence continuous in z. Hence, for any fixed φ, ζ continuously maps the compact and convex set $[0,1]^2$ into itself, and we can apply Brouwer's fixed point theorem to the map $\zeta(z(\mu,\varphi))$. □

Fix a root $\mu^*(\varphi) \in \Upsilon^*(\varphi)$, and let $z^*(\varphi) \equiv \varphi\mu_1^*(\varphi) + (1-\varphi)\mu_2^*(\varphi)$ be the population state played at φ. Then, the Jacobi $C_{\mu^*}(\varphi) \equiv \partial_\mu g(\mu^*(\varphi),\varphi)$ for $\varphi \in [0,1]$ is given by

$$C_{\mu^*}(\varphi) = \begin{pmatrix} \zeta_1'(z^*(\varphi))\,\varphi - 1 & \zeta_1'(z^*(\varphi))(1-\varphi) \\ \zeta_2'(z^*(\varphi))\,\varphi & \zeta_2'(z^*(\varphi))(1-\varphi) - 1 \end{pmatrix}$$

Later, we will be mainly interested in stable roots at the belief-monomorphic state $\varphi = 1$ where the eigenvalues of $C_{\mu^*}(\varphi)$ all have strictly negative real parts. The next lemma characterizes such stable roots. Let $B_\delta(x) \equiv \{y \in [0,1] : \|x-y\| < \delta\}$ for $x \in [0,1]$. Then we have:

Lemma 2. *Fix $\tilde\varphi = 1$ a profile $\{F_1,F_2\} \in \mathscr{F} \times \mathscr{F}$, and take any root $\mu^*(\varphi)$ with $\mu^*(\tilde\varphi) \in \Upsilon^*(\tilde\varphi)$. If $C_{\mu^*}(\varphi)$ has strictly negative real eigenvalues, then there exists $\delta > 0$ such that*

(a) *$\mu^*(\varphi)$ is differentiable $\forall \varphi \in B_\delta(\tilde\varphi)$, and*
(b) *$C^*(\varphi)$ has eigenvalues with negative real parts bounded away from zero $\forall \varphi \in B_\delta(\tilde\varphi)$.*

The proof is found in Appendix A.1. For later purpose we note (i) that the first element $\mu_1^*(\tilde\varphi)$ of $\mu^*(\tilde\varphi)$ only depends on ζ_1, and (ii) that the eigenvalues of $C_{\mu^*}(\tilde\varphi)$ only depend on the reaction function ζ_1 of agents in subpopulation 1. The latter follows as neither the trace $\mathrm{tr}(C_{\mu^*}(\tilde\varphi))$ nor the determinant $\det(C_{\mu^*}(\tilde\varphi))$ depend on ζ_2, and the two eigenvalues λ_1,λ_2 satisfy $\lambda_1+\lambda_2 = \mathrm{tr}(C_{\mu^*}(\tilde\varphi))$ and $\lambda_1\lambda_2 = \det(C_{\mu^*}(\tilde\varphi))$. Hence, the stability of any $\mu^*(\tilde\varphi)$ only depends on $F_1 \in \mathscr{F}$. The converse holds, of course, for the roots $\mu^*(\varphi)$ at $\varphi = 0$.

So far, we have not specified what particular belief distribution $F \in \mathscr{F}$ we might expect agents to draw from. This indeterminacy of the belief distribution translates in an indeterminacy of the best response rest point. The reason for this is that the rest point of the best response node largely depends on the particular belief distributions that agents draw from, whereas the underlying payoff structure only has a very coarse influence. The next example makes this more explicit:

Example 2. Let $\mathscr{F}$ be the family of belief distributions generated by random sampling as described in Example 1, and consider a simple hawk-dove game with payoffs $U_1(z) = -(1+\epsilon)z$ and $U_2 = -(1-z)$ with $\epsilon > 0$. This game has a unique Nash equilibrium $z^* = 1/(2+\epsilon) < 1/2$. That is, an agent with a belief $\hat{z} > z^*$ chooses strategy 2, whereas an agent with a belief $\hat{z} \leq z^*$ chooses strategy 1. Let the arrival rate of revision opportunities be $\lambda = 1$, assume for the sake of clarity that all agents of the population draw from the same belief distribution, and that hence we can take the population state z itself to be subject to the best response dynamics. Take first the case of agents sampling exactly $M = 1$ other agent when revising their beliefs. Under such a sampling regime, an agent will play strategy 1 iff the sample contains a strategy 2 encounter. This happens with probability $1-z$. The dynamics are then given by $\dot{z} = 1-2z$. For $M = 1$ the rest point is given by $\tilde{z} = 1/2 \neq z^*$. For $M = 2$, an agent chooses strategy 1 only when he has two strategy 2 agents in his sample, so the dynamics are given by $\dot{z} = (1-z)^2 - z$, and hence it holds at the rest-point $\tilde{z}$ that $\tilde{z} = (1-\tilde{z})^2$. Obviously, the $\tilde{z}$ that satisfies this conditions is neither equal to the mixed Nash equilibrium strategy share z^*, nor is it equal to the rest-point under $M = 1$. The rest point changes when we change the sampling procedure and hence change the belief distribution.

Consequently, in order that the rest point of the best response dynamics is not totally arbitrary, we need to introduce a selection device between belief distribution. This is done next.

2.2 Evolution

In this section, we describe the dynamics of the sub-population share φ. We assume the dynamics to be evolutionary in the sense that the sub-population share of those agents receiving a higher average sub-population utility increases, whereas the share of those agents receiving the lower average sub-population utility shrinks. To model the growth rate of φ, we adapt the idea of the replicator dynamics (cf. Taylor & Jonker 1978). The average sub-population utility in sub-population $q = 1, 2$ is given by

$$\overline{U}_q(\mu, \varphi) \equiv \mu_q U_1(z(\mu, \varphi)) + (1-\mu_q) U_2(z(\mu, \varphi))$$

The time of evolution is denoted by s. We write the evolution node as

$$\frac{\mathrm{d}\varphi(s)}{\mathrm{d}s} = \varphi(s)\,(1-\varphi(s))\left(\overline{U}_1\,(\mu(s), \varphi(s)) - \overline{U}_2\,(\mu(s), \varphi(s))\right)$$

or, equivalently as

$$\frac{\mathrm{d}\varphi(s)}{\mathrm{d}s} = \varphi(s)\,(1-\varphi(s))\,(\mu_1(s)-\mu_2(s))\,(U_1\,(z(\mu(s),\varphi(s)))-U_2\,(z(\mu(s),\varphi(s))))$$
$$\equiv h\,(\mu(s),\varphi(s))$$

with $\varphi \in [0,1]$. This expression has a natural interpretation: As long as we are not in a belief-monorphic state with $\varphi \in \{0,1\}$, the share φ of belief distribution F_1 grows if either strategy 1 yields a higher utility than strategy 2 and the sub-population drawing beliefs from F_1 has currently a higher fraction of agents choosing strategy 1, or if strategy 2 yields a higher utility than strategy 1 and the sub-population drawing beliefs from F_1 has currently a higher fraction of agents choosing strategy 2.

2.3 Time-scales

The time-scale of the evolutionary dynamics (s) differs from the time-scale of the best-response dynamics (t). In particular, we assume that evolution runs on a slower time-scale than the best response dynamics, that is, we have $s/t = \epsilon$ with $0 < \epsilon \leq 1$. The parameter ϵ stands for the ratio of the speed of evolution over the speed of the best response dynamics.

In the following, we analyze the case of small, but positive ϵ values. For small ϵ values, the replicator dynamics determining the belief distribution share φ is much slower compared to the best response dynamics determining the strategy distributions μ. In the limit of $\epsilon \to 0$, the best response dynamics becomes infinitely fast, that is, behavior adapts instantaneously to changes in the belief distribution shares.

We use the dot-notation as in $\dot{\mu}$ to denote derivatives with respect to the slow time s. We have

$$\dot{\mu}_q = \frac{\mathrm{d}\mu_q(t)}{\mathrm{d}s} = \frac{\mathrm{d}\mu_q(t)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}s} = \frac{\mathrm{d}\mu_q(t)}{\mathrm{d}t}\frac{1}{\epsilon} = \frac{1}{\epsilon}\delta\left(\zeta_q(z(\mu(t),\varphi(t))) - \mu_q(t)\right),$$

and hence we can rewrite the joint dynamics of the best response dynamics and the replicator dynamics as

$$\begin{array}{rcll} \epsilon\dot{\mu} & = & g(\mu,\varphi), & \mu \in [0,1]^2 \\ \dot{\varphi} & = & h(\mu,\varphi), & \varphi \in [0,1] \end{array} \tag{4}$$

The dynamical system (4) is at the center of interest in this paper. Resorting to the terminology of slow-fast systems, we call $\varphi \in [0,1]$ the slow variable, and the components of $\mu \in [0,1]^2$ the fast variables. Alternatively, the latter is called the best response node, whereas the former is called the evolution node.

Note that the two nodes are uncoupled when $\epsilon = 0$: The fast node variables μ are at rest for any φ, and the slow node is thus dependent on φ alone (once through the rest-point of the replicator dynamics determined by φ and once through φ directly). In the following analysis, we derive conditions under which the behavior of (μ, φ) at $\epsilon = 0$ can be taken as a good approximation for the behavior of (μ, φ) for small $\epsilon > 0$.

3 Stability

We restrict attention to belief-monomorphic rest points $(\tilde{\mu}, \tilde{\varphi})$ of system (4) with $\tilde{\varphi} = 1$ and with the Jacobi $\partial_\mu g(\tilde{\mu}, \tilde{\varphi})$ having only strictly negative real eigenvalues. That is, we look at a situation where we have a resident population of agents drawing revised beliefs from $F_1 \in \mathscr{F}$ yielding stable root $\mu^*(\tilde{\varphi}) = \tilde{\mu}$, and a candidate mutant distribution $F_2 \in \mathscr{F}$. For our stability analysis, we fix such a pair $F_1, F_2 \in \mathscr{F}$ of belief distributions, consider a perturbation (μ_0, φ_0) of the rest point $(\tilde{\mu}, \tilde{\varphi})$, and then look at the resulting trajectories (μ_t, φ_t) of the slow-fast system (4).

In order to describe the behavior of system (4) after the perturbation (μ_0, φ_0), we resort to results which are originally due to Tikhonov, and are restated in Kokotovic (1984). By Lemma 2, we know that there is a $\delta > 0$ such that the root $\mu^*(\varphi) \in \Upsilon^*(\varphi)$ with $\mu^*(\tilde{\varphi}) = \tilde{\mu}$ is differentiable $\forall \varphi \in B_\delta(\tilde{\varphi})$. Hence, the following reduced system is well defined on $B_\delta(\tilde{\varphi})$:

$$\dot{\overline{\varphi}} = h(\mu^*(\overline{\varphi}), \overline{\varphi}), \quad \overline{\varphi} \in B_\delta(\tilde{\varphi}) \tag{5}$$

The reduced system (5) describes the evolution of the belief distribution shares given that the best-response process adapts infinitely fast to changes in the shares of belief distributions. For the following Theorem, we collect the results, and in particular, Theorem 2.1 in Kokotovic (1984). Consider the slow-fast system (4) with initial values $(\mu_0, \varphi_0) \in [0,1]^3$, and the reduced system (5) with initial value $\overline{\varphi}_0 = \varphi_0$, and solution trajectory $\overline{\varphi}_t$.

Theorem 1. *Assume that the eigenvalues of the Jacobi $C^*(\overline{\varphi}_t) = \partial_\mu g(\mu^*(\overline{\varphi}_t), \overline{\varphi}_t)$ all have real parts that are strictly below some fixed negative number $\forall t \in [0, T]$, $T > 0$. Then, $\exists \omega > 0$ such that if $\|\mu_0 - \mu^*(\varphi_0)\| \leq \omega$ there exists $t_0 > 0$, $t_0 \leq T$, with $t_0 \to 0$ as $\epsilon \to 0$, such that*

$$\|\varphi_t - \overline{\varphi}_t\| = O(\epsilon), \quad \forall t \in [0, T] \tag{6}$$

$$\|\mu_t - \mu^*(\overline{\varphi}_t)\| = O(\epsilon), \quad \forall t \in [t_0, T] \tag{7}$$

Given Theorem 1 applies, equation (6) states that the trajectory $\overline{\varphi}_t$ of the reduced system (5) is a good approximation for the behavior of the trajectory φ_t of the full system (4) as long as $t \in [0, T]$, given we choose μ_0 sufficiently close to $\mu^*(\varphi_0)$. Similarly, but only after some non-negative time t_0, the trajectory μ_t of the full system can be approximated by $\mu^*(\overline{\varphi}_t)$ until $t = T$. This is stated with (7).

Theorem 1 is convenient for our analysis of the stability of belief distribution as it suggests to consider initial conditions (μ_0, φ_0) that can be expressed in terms of φ_0 alone: Choose some φ_0 in the domain $B_\delta(\tilde{\varphi})$ where the reduced system (5) is defined, and pick μ_0 sufficiently close to $\mu^*(\varphi_0)$ such that $\|\mu_0 - \mu^*(\varphi_0)\| \leq \omega$ with ω as given in Theorem 1. Hence, when we talk about stability in the following, we always only make explicit reference to φ_0, but nevertheless always implicitly assume μ_0 to lie sufficiently close to $\mu^*(\varphi_0)$. This conception of a perturbation is agnostic about how the resident population reacts to a sudden small change in the sub-population share φ. We merely assume that the strategy shares are minimally and arbitrarily perturbed when mutants intrude.

3.1 Replicator Dynamics Stable Belief Distributions

The candidate mutant distribution F_2 is assumed to be different from F_1 in a specific sense: Letting again $B_\delta(x) = \{y \in [0,1] : \|x - y\| < \delta\}$ for $x \in [0,1]$, we define:

Definition 1 (z-different belief distribution). F_1 is z-different from F_2, if $\exists \omega > 0$ such that $\|\zeta_1(z') - \zeta_2(z')\| \neq 0$, $\forall z' \in B_\omega(z)$.

As we have differentiable reaction functions, a necessary and sufficient conditions for two distributions F_1 and F_2 to be z-different is that $\|\zeta_1(z) - \zeta_2(z)\| \neq 0$ holds. The notion of z-different belief distributions is important for our stability criterion, as it excludes cases in which the dynamics get "stuck" because the trajectory of (μ_t, φ_t) at some time t passes $(\mu^*(\varphi_t), \varphi_t)$ with a population state $z_t = z(\mu^*(\varphi_t), \varphi_t)$ at which both the resident and the intruding belief distributions yield the same strategy distributions. In particular, if for $\zeta_1(\tilde{z}) = \zeta_2(\tilde{z})$ initial conditions $\varphi_0 \neq \tilde{\varphi}$ and $\mu_0 = (\tilde{z}, \tilde{z})$ are chosen[3], then we have $h(\mu_0, \varphi_0) = g(\mu_0, \varphi_0)$, and the system is at rest right after the perturbation.

Having ruled out such cases, we consider a belief distribution $F_1 \in \mathscr{F}$, and let $\tilde{z} = z(\tilde{\mu}, \tilde{\varphi})$ be the belief-monomorphic population state. As at $\tilde{\varphi} = 1$, the population state $\tilde{z}$ under any $\mu^*(\tilde{\varphi}) \in \Upsilon^*(\tilde{\varphi})$ depends solely on F_1, Theorem 1 motivates the following notion of stability:

[3]This is consistent with our notion of perturbation if φ_0 is not too far from $\tilde{\varphi}$.

Definition 2 (Replicator Dynamics Stable Belief Distribution)**.** The belief distribution $F_1 \in \mathscr{F}$ is replicator dynamics stable at $\tilde{z}$, if for any belief distribution $F_2 \in \mathscr{F}$ that is $\tilde{z}$-different from F_1, and for all $\epsilon > 0$ small enough it holds that

(a) for every $\Sigma > 0$, there exists a $\Delta(\Sigma) > 0$ such that if $\|\varphi_0 - \tilde{\varphi}\| \leq \Delta(\Sigma)$, then $\|\varphi_t - \tilde{\varphi}\| \leq \Sigma + O(\epsilon)$ for all $t \geq 0$, and
(b) there exists a $\Delta > 0$ such that if $\|\varphi_0 - \tilde{\varphi}\| \leq \Delta$, then $\lim_{t\to\infty} \varphi_t \in B_{O(\epsilon)}(\tilde{\varphi})$

This definition establishes a belief distribution as stable under the replicator dynamics at population state $\tilde{z}$, if the intrusion of a small share of any belief distribution that yields a differing reaction function at $\tilde{z}$ cannot be successful in the sense that, in the long run, the mutants – beyond possibly a small fraction that vanishes in the speed of the best-response dynamics – are crowded out again. The qualification of stability at a given state $\tilde{z}$ is important, as the belief distribution F_1 might yield multiple roots $\mu^*(\tilde{\varphi})$.

With $\epsilon = 0$, Definition 2 reads like a definition of asymptotic stability for $\tilde{\varphi}$. This is no surprise as with $\epsilon = 0$ the evolution and the best response node are uncoupled, and the behavior of φ_t is equivalent to the behavior of the trajectory $\overline{\varphi}_t$ of the reduced system. The relation to the reduced system holds for $\epsilon > 0$, too, as the following main result of this section shows. As at $\tilde{\varphi} = 1$, the eigenvalues of $C_{\mu^*}(\tilde{\varphi})$ under any $\mu^*(\tilde{\varphi}) \in \Upsilon^*(\tilde{\varphi})$ depend solely on F_1, it follows from combining Theorem 1 and Lemma 2 that:

Theorem 2. *Let $\mathscr{F}$ comply with Assumptions 1 (Differentiability) and 2 (Richness). Consider a belief distribution $F_1 \in \mathscr{F}$ that yields $\mu^*(\tilde{\varphi})$ with $\mu_1^*(\tilde{\varphi}) = \tilde{z}$ and with the eigenvalues of $C_{\mu^*}(\tilde{\varphi})$ all having strictly negative real parts. Then, F_1 is replicator dynamics stable at $\tilde{z}$ if and only if $\tilde{\varphi}$ is an asymptotically stable rest point of the reduced system (5) for any belief distribution $F_2 \in \mathscr{F}$ that is $\tilde{z}$-different from F_1.*

Proof. We first show the if-part of the statement. If $\tilde{\varphi}$ is asymptotically stable on the reduced system, then we can always bound the neighborhood through which the trajectory $\overline{\varphi}_t$ passes by choosing $\overline{\varphi}_0$ close enough to $\tilde{\varphi}$. In particular, we can choose $\overline{\varphi}_0$ such that $\overline{\varphi}_t$ does not leave the neighborhood in which $C^*(\varphi)$ has eigenvalues with real part strictly bounded away from zero. Hence, $T = \infty$ and we have $\|\varphi_t - \overline{\varphi}_t\| = O(\epsilon)\ \forall t \geq 0$. Now, recall the definition of asymptotic stability for $\tilde{\varphi}$ on the reduced system $\overline{h}$: A rest-point $\tilde{\varphi}$ of $\overline{h}$ is asymptotically stable (i) if for every $\gamma > 0$ there exists a $\delta = \delta(\gamma) > 0$ such that if $\|\overline{\varphi}_0 - \tilde{\varphi}\| < \delta$, then $\|\overline{\varphi}_t - \tilde{\varphi}\| < \gamma$, $\forall t \geq 0$, and (ii) if there exists $\delta > 0$ such that if $\|\overline{\varphi}_0 - \tilde{\varphi}\| < \delta$, then $\lim_{t\to\infty} \|\overline{\varphi}_t - \tilde{\varphi}\| = 0$. Together with the result that $\|\varphi_t - \overline{\varphi}_t\| = O(\epsilon)\ \forall t \geq 0$, the statement then follows.

To show the only if part of the statement, first suppose that $\tilde{\varphi}$ is not neutrally stable on the reduced system. Then there exists an $\overline{\Sigma}$ such that for all $\Sigma \leq \overline{\Sigma}$, there is no $\Delta > 0$ such that $\|\overline{\varphi}_0 - \tilde{\varphi}\| \leq \Delta$ implies $\|\overline{\varphi}_t - \tilde{\varphi}\| \leq \Sigma$, $\forall t > 0$. Nevertheless, there is a $T > 0$, such that $C^*(\overline{\varphi}_t) = \partial_\mu g(\mu^*(\overline{\varphi}_t), \overline{\varphi}_t)$ all have real parts that are strictly below some fixed negative number $\forall t \in [0, T]$, and hence $\|\varphi_t - \overline{\varphi}_t\| = O(\epsilon)$ holds for those $t \in [0, T]$. Consequently, there is an $\epsilon > 0$ small enough such that there exists an $\overline{\Sigma}$ such that for all $\Sigma \leq \overline{\Sigma}$, there is no $\Delta > 0$ such that $\|\varphi_0 - \tilde{\varphi}\| \leq \Delta$ implies $\|\varphi_t - \tilde{\varphi}\| \leq \Sigma + O(\epsilon)$. Second, suppose that $\tilde{\varphi}$ is neutrally stable but not asymptotically stable on the reduced system. Again, if $\tilde{\varphi}$ is neutrally stable on the reduced system, then we can always bound the neighborhood through which the trajectory $\overline{\varphi}_t$ passes by choosing $\overline{\varphi}_0$ close enough to $\tilde{\varphi}$. In particular, we can choose $\overline{\varphi}_0$ such that $\overline{\varphi}_t$ does not leave the neighborhood in which $C^*(\varphi)$ has eigenvalues with real part strictly bounded away from zero. Hence, it holds that $T = \infty$, and we have $\|\varphi_t - \overline{\varphi}_t\| = O(\epsilon)$ $\forall t \geq 0$. As $\tilde{\varphi}$ is not asymptotically stable, we have $\lim_{t\to\infty} \overline{\varphi}_t \neq \tilde{\varphi}$. Hence, there is an $\epsilon > 0$ small enough such that $\lim_{t\to\infty} \varphi_t \notin B_{O(\epsilon)}(\tilde{\varphi})$. As we can repeat these two arguments for any $\overline{\varphi}_0$, we have the statement. □

The result is useful as it directs our attention to the reduced system when we want to establish replicator dynamic stability of F_1 at some $\tilde{z}$. The reduced system has the same number of dimensions as the evolutionary node. This reduction in dimensions will become particularly useful when we extend the setting to a two-population model with two strategies in Section 5. Already with this simple extension, the system comprising both the best response and the replicator dynamics turns out to be analytically untractable, whereas the reduced system can still be analyzed.

3.2 Relation to Replicator Dynamics Stable Strategies

We now analyze how the belief distributions that are stable under the replicator dynamics at some population state $\tilde{z}$ relate to the stable strategy distribution z that obtains under the replicator dynamics in the traditional evolutionary game theory approach (cf. Weibull 1997). In particular, we are interested in how the relation is dependent on the assumptions that we make on $\mathscr{F}$. The section yields two results that we use later in the applications. The traditional replicator dynamics on the strategy share z is given by

$$\dot{z} = z(1-z)[U_1(z) - U_2(z)] \equiv z(1-z)k(z) \tag{8}$$

A rest point $\tilde{z}$ of this dynamics is asymptotically stable iff there exists a neighborhood around $\tilde{z}$ such that it holds that $(z - \tilde{z})k(z) < 0$ for all $z \neq \tilde{z}$ in that neighborhood. For $\tilde{z} = 1$, this is equivalent to the condition that $U_1(\tilde{z}) > U_2(\tilde{z})$, and for $\tilde{z} = 0$, this is equivalent to the condition that $U_1(\tilde{z}) < U_2(\tilde{z})$. If k is differentiable in z – as it is in the case of normal form games – then for interior $\tilde{z}$, this is equivalent to the condition that $k'(\tilde{z}) < 0$ (cf. Weibull 1997). If $\tilde{z}$ is asymptotically stable in this sense, we call the strategy distribution $\tilde{z}$ replicator dynamics stable.

We are interested in how the dynamics (8) relate to the slow-fast dynamics as described in the last section. The reduced system (5) can be written as

$$\dot{\varphi} = \varphi(1-\varphi)(\mu_1^*(\varphi) - \mu_2^*(\varphi))[U_1(z^*(\varphi)) - U_2(z^*(\varphi))], \quad \varphi \in B_\delta(\tilde{\varphi}) \tag{9}$$

Comparing the replicator dynamics (8) abd the reduced system (9), we arrive at the following first result:

Proposition 1. *Let $\mathscr{F}$ comply with Assumptions 1 (Differentiability) and 2 (Richness). Consider a belief distribution $F_1 \in \mathscr{F}$ that yields $\mu^*(\tilde{\varphi})$ with $\mu_1^*(\tilde{\varphi}) = \tilde{z}$ and with the eigenvalues of $C_{\mu^*}(\tilde{\varphi})$ all having strictly negative real parts. Then, F_1 is replicator dynamics stable at $\tilde{z}$ iff $\tilde{z}$ is an asymptotically stable rest point of the replicator dynamics* (8).

The proof is found in Appendix A.2. We will use this result in Section 4 where we look at specific belief distributions that are brought about by random sampling. Furthermore, in Section 5, we show that the equivalence result does not hold for two-population games.

We next present a third assumption about the family $\mathscr{F}$ of belief distributions under which in games of the Hawk-Dove variety (1) the root $\mu^*(\varphi)$ is unique, and (2) the eigenvalues of $C_{\mu^*}(\tilde{\varphi})$ all have strictly negative real parts. The assumption concerns the informativeness of the distribution, and we call a family $\mathscr{F}$ that complies with it regular:

Assumption 3 (Regularity)**.** *It holds for any $F \in \mathscr{F}$ that for two strategy shares $u, v \in [0,1]$ with $u > v$, it follows that $F(x|u) \leq F(x|v)$, $\forall x \in [0,1]$, where the inequality is strict for at least some $x \in [0,1]$.*

A belief distribution of a regular family $\mathscr{F}$ is informative in the following specific sense: For all conditional belief distributions that are available to the agents, the belief distribution conditional on an actual strategy 1 share $u \in [0,1]$ stochastically dominates the distribution conditional on a share $v < u$. That is, high signal become more likely as z grows.

The implications for the shape of the reaction functions ζ_q, $q = 1,2$ are straight-forward: For Coordination Games with $a, b > 0$, we have $\zeta_q'(z) \geq 0$, $\forall F \in \mathscr{F}$, and for Hawk-Dove games, we have $\zeta_q'(z) \leq 0$, $\forall F \in \mathscr{F}$. In the former case, a higher share of strategy 1 players in the population goes along with higher probabilities that the signal is such that the updating agent chooses strategy 1, too. In the latter case, the reverse holds true.

Note that from $\zeta_q'(z) \leq 0$ it follows immediately that $\mu^*(\varphi)$ is unique for any $\varphi \in [0,1]$ in the Hawk-Dove game. Further, we get the following equivalence result for games of the Hawk-Dove variety as a Corollary to Proposition 1:

Corollary 1. *Suppose $a, b < 0$, and let $\mathscr{F}$ comply with Assumptions 1 (Differentiability), 2 (Richness) and 3 (Regularity). Then any belief distribution $F_1 \in \mathscr{F}$ that yields $\mu^*(\tilde{\varphi})$ with $\mu_1^*(\tilde{\varphi}) = \tilde{z}$ is replicator dynamics stable at $\tilde{z}$ iff $\tilde{z}$ is an asymptotically stable rest point of the replicator dynamics* (8).

Proof. By Proposition 1, we merely need to show that the assumption of regularity implies that $C^*(\tilde{\varphi})$ has eigenvalues whose real parts are strictly below zero. If $a, b < 0$, then under regular belief distributions, we have $\zeta'(z) \leq 0$ for any $F \in \mathscr{F}$. From the proof of Proposition 1 in Appendix A.2, it follows that $\zeta'(z) \leq 0$ implies that $C^*(\tilde{\varphi})$ has eigenvalues whose real parts are strictly below zero. Hence, the result follows. □

As only the mixed Nash equilibrium is replicator dynamics stable, it thus follows for games of the Hawk-Dove variety that there exists a unique population state that can be supported by a replicator dynamics stable belief distribution. We will make use of this result in Section 4.2 where we look at random sampling in games of the Hawk-Dove variety.

4 Application I: Generalized Random Sampling

We now look at the particular family of belief distributions that results from random sampling as introduced in Example 1, and briefly discussed in Example 2. The sampling model has some intuitive appeal as it provides a natural description of how agents gather and process information. The idea is used, for example, in the word-of-mouth learning literature (Ellison & Fudenberg 1993, Ellison & Fudenberg 1995, Banerjee & Fudenberg 2004), or in models of consumer choice behavior (e.g. Spiegler 2006a, Spiegler 2006b). The best response dynamics that emerge under random sampling are extensively analyzed in Sandholm (2001*a*), and Oyama, Sandholm & Tercieux (forthcoming). In contrast to the existing research that takes an unbiased sampling procedure as given, we extend the basic notion of random sampling to generalized sampling that allows for biased sampling probabilities. This will allow us to ask under what circumstances unbiased sampling induces belief distributions that are replicator dynamics stable in our sense.

Let $\mathscr{C}^1_M([0,1],[0,1])$ denote the space of continuously differentiable, monotonically increasing functions from $[0,1]$ to $[0,1]$. Adapting the definition of unbiased sampling given in Example 1, we define generalized sampling as follows:

Definition 3 (Generalized Sampling)**.** A generalized sampling procedure is described by (M,p) where $M \in \mathbb{N}_{++}$ is the sample size, and $p \in \mathscr{C}^1_M([0,1],[0,1])$ returns the sampling probability $p(z)$ of strategy 1 at a given strategy distribution $z \in [0,1]$.

Definition 4 (Locally Unbiased Sampling Rule)**.** A sampling rule (M,p) is called locally unbiased at $z \in [0,1]$ if $p(z) = z$.

A sampling rule with $p(z) = z$, $\forall z \in [0,1]$, is accordingly called globally unbiased. Let $m \leq M$ be the number of strategy 1 agents in the sample. Then, the empirical mean $\hat{z} = n/M$ is the resulting belief about the strategy 1 share in the population. The family of belief distributions generated by generalized random sampling is denoted by $\mathscr{F}_S$. Beliefs $\hat{z}$ under random sampling are binomially distributed with cumulative distribution function

$$F(x|z) = \sum_{i=1}^{\lfloor M\cdot x \rfloor} \binom{M}{i} p(z)^i (1-p(z))^{M-i}$$

Note that we can rewrite this alternatively as

$$F(x|z) = I_{1-p(z)}(M - \lfloor M \cdot x \rfloor, \lfloor M \cdot x \rfloor + 1)$$

where $I_x(\alpha, \beta)$ denotes the cumulative distribution function of the beta distribution with coefficients α, β. We have used the fact that the value $F_X(x)$ of the cumulative distribution function F_X of a binomial variable $X \sim B(n,p)$ evaluated at x is equal to the value $F_Y(1-p)$ of the distribution function F_Y of a beta-distributed variable $Y \sim Beta(n-x, x+1)$ evaluated at $1-p$ (cf. Olver, Lozier, Boisvert & Clark 2010).

Clearly, $F(x|z)$ is differentiable in z, as $I_x(\alpha,\beta)$ is differentiable in x, and $p(z)$ is differentiable in z. Furthermore, the family $\mathscr{F}_S$ of distributions generated by generalized sampling rules is rich, and its members $F \in \mathscr{F}_S$ are regular, as the next lemma asserts.

Lemma 3. *Let $F(x|z) \in \mathscr{F}_S$. Then:*

(a) For any $u, v, z \in [0,1]$, $\exists (M,p)$ such that $u = F(v|z)$.
(b) For any $u, v, x \in [0,1]$, $u > v$, it follows that $F(x|u) \leq F(x|v)$.

Proof. (a): Write $u = F(v|z) = I_{1-p(z)}(M - \lfloor Mv \rfloor, \lfloor Mv \rfloor + 1)$. Fix $v \in [0,1]$, $M \in \mathbb{N}_{++}$. As we have $I_0(.,.) = 0$, $I_1(.,.) = 1$ and $0 < \partial I_x/\partial x < \infty$, we can choose $p(z) \in [0,1]$ such that $u = I_{1-p(u)}(.,.)$. As we can do this for any $v \in [0,1]$, and $M \in \mathbb{N}_{++}$, we have the claim. (b): This follows directly from the fact that $p(z)$ is increasing in z, and $I_{1-p(z)}(.,.)$ is decreasing in $p(z)$. □

With these preliminary comments, we now look at the two varieties $a, b > 0$, and $a, b < 0$ separately.

4.1 Coordination Games

We first analyze the case of $a, b > 0$. Let $\bar{z} \equiv b/(a+b)$. Assuming that the agents plays strategy 1 whenever strategy 1 is a weak best reply to their belief $\hat{z}$, we have for agents of sub-population q using a sampling rule (M, p):

$$\zeta_q(z) = \sum_{i=\lceil M\cdot\bar{z}\rceil}^{M} \binom{M}{i} p(z)^i (1-p(z))^{M-i} \tag{10}$$

As the two pure strategy equilibria $z^* \in \{0,1\}$ are the only strategy distributions that are stable under the replicator dynamics, it follows immediately from Proposition 1 that any $F_1 \in \mathscr{F}_{\mathscr{S}}$ that yields $\mu^*(\tilde{\varphi})$ with $\mu_1^*(\tilde{\varphi}) = \tilde{z}$ and $C_{\mu^*}(\tilde{\varphi})$ only having eigenvalues with strictly negative real parts is replicator dynamics stable at $\tilde{z}$ iff $\tilde{z} \in \{0,1\}$. Focusing thus on $\tilde{z} \in \{0,1\}$, we arrive at:

Proposition 2. *For any $\tilde{z} \in \{0,1\}$, $\exists (M,p)$ that generates a belief distribution which is replicator dynamics stable at $\tilde{z}$.*

From (10), it immediately follows that any (M,p) that generates a belief distribution which is replicator dynamics stable at $\tilde{z} \in \{0,1\}$ is locally unbiased.

4.2 Hawk-Dove Games

The more interesting case is $a, b < 0$. Again, let $\bar{z} \equiv b/(a+b)$. As the belief distributions are regular, we can apply Corollary 1 to conclude that the mixed strategy equilibrium $\bar{z}$ is the unique strategy distribution that a belief distribution in $\mathscr{F}$ can be replicator dynamics stable at. A sampling procedure (M,p) that leads to a belief distribution which yields $\bar{z}$ as a rest point of the best response node is henceforth called stable:

Definition 5 (Stable Sampling). A sampling procedure (M,p) is called stable for $\bar{z}$, if the sample size M and the sampling probability p are such that

$$\bar{z} = I_{1-p(\bar{z})}(M - \lfloor M\bar{z} \rfloor, \lfloor M\bar{z} \rfloor + 1)$$

Stable sampling is not always unbiased. The following proposition shows that being locally biased at $\bar{z}$ is indeed generic to the set of stable sampling procedures. We think of generic local biasedness as follows: Let the payoffs $a, b < 0$ be drawn from some continuous distribution such that the equilibrium $\bar{z}$ itself can be represented as a draw from a random variable $\mathscr{Z} \in (0,1)$ with some continuous distribution function $Z : (0,1) \to [0,1]$. Suppose that the population has been cast into the game long ago, and evolution has selected a belief distribution $F_1 \in \mathscr{F}_S$ that yields $\mu^*(\tilde{\varphi})$ with $\mu_1^*(\tilde{\varphi}) = \bar{z}$. Let

$$p_M^*(\bar{z}) = \{x \in [0,1] : \bar{z} = I_{1-x}(M - \lfloor M\bar{z} \rfloor, \lfloor M\bar{z} \rfloor + 1)\}$$

be the sampling probability for population state $\bar{z}$ that renders a sampling procedure with sample size M stable at $\bar{z}$. We have the following result:

Proposition 3. *Let $\bar{z}$ be a draw from a random variable $\mathscr{Z} \in (0,1)$ with continuous distribution function $Z : (0,1) \to [0,1]$. Then, for any bound $\overline{M} \in \mathbb{N}_{++}$ on the sample size, the set $M_0(\bar{z}) \equiv \{M \leq \overline{M} : p_M^*(\bar{z}) = \bar{z}\}$ is empty with probability one.*

The proof is in Appendix A.4. The proposition states that the probability of an equilibrium $\bar{z}$ that can be supported by a replicator dynamics belief distribution which is generated by an unbiased sampling procedure is zero. There is, with probability one, no unbiased sampling rule that is stable. The reason for this phenomenon is relatively simple: Sampling produces belief distributions $F(x|z)$ with discontinuities in x. Therefore, there are equilibria $\bar{z} \in (0,1)$ that are not stable under random beliefs generated by unbiased sampling rules. In fact, as Proposition 3 shows, the set of such equilibria has measure 1. The sampling probability $p(\bar{z})$ is then needed to correct the induced belief distribution such that the reaction functions yield $\bar{z}$ as a steady state of the best response node.

5 Application II: Matching Pennies

This section extends the framework to asymmetric 2×2-normal form games, and looks at varieties of the Matching Pennies game. Matching Pennies is of particular interest since the unique Nash equilibrium of the game is not stable under the replicator dynamics (cf. Ritzberger & Weibull 1995). Furthermore, it is not generically selected under the best-response dynamics with stochastic beliefs, either. Combining the two approaches, however, puts the unique mixed Nash equilibrium back into focus. In particular, we can show that, if we assume regularity of the family $\mathscr{F}$, then the unique mixed strategy equilibrium is the unique strategy profile at which a belief distribution $F \in \mathscr{F}$ can be replicator dynamics stable. That is, the equivalence that we observe in symmetric 2×2-normal form games (cf. Proposition 1) does not carry over to asymmetric 2×2 matching pennies games.

5.1 The Setup

The game has two populations $p = A, B$ whose agents are randomly pairwise matched across populations in every round. Each population consists of two sub-populations $q = 1, 2$. In the following, the index $-p$ always refers to population p's opponent population. As the strategy space is 2-dimensional for both populations, we now write z^p for the share of agents in population $p = A, B$ playing strategy 1, and let $z = (z^A, z^B)$. The expected utility $U_i^p(z^{-p})$ in a match to an agent of population p playing strategy $i = 1, 2$ at opponent population state z^{-p} is given by $U_1^p(z^{-p}) = a_p z^{-p}$, and $U_2^p(z^{-p}) = b_p(1 - z^{-p})$, with $a_p, b_p \in \mathbb{R}$ satisfying $a_A, b_A < 0$, and $a_B, b_B > 0$. That is, the two populations have distinct roles in this game: Agents from population B are the *matchers* preferring the other agent to chose the same strategy as they do, whereas agents from population A are the *mismatchers* preferring the opponent to play the other strategy. Consequently, there is no Nash equilibrium in pure strategies. The unique mixed strategy equilibrium is given by $z^* = \left(z^{A^*}, z^{B^*}\right)$ with $z^{p^*} = b_{-p}/(a_{-p} + b_{-p})$, $p = A, B$.

Agents of Population p that revise their beliefs at a population state z^{-p} draw a belief $\hat{z}^{-p}$ from continuous distribution $F(.|z^{-p}) \in \mathscr{F}$ that is differentiable in z^{-p} and has density $f(.|z^{-p}) \geq 0$ on $[0, 1]$. $\mathscr{F}$ is the same for both populations, and we denote the belief distribution of the agents of population p, sub-population q, by $F_q^p \in \mathscr{F}$. We take Assumptions 1 (Differentiability), 2 (Richness) and 3 (Regularity) to hold for $\mathscr{F}$.

Agents play strategy 1 whenever this is a best reply to their belief about the state of the opponent population. Then, for $q = 1, 2$, the reaction functions are given by

$$\zeta_q^A\left(z^B\right) = F_q^A\left(z^{B^*} \middle| z^B\right)$$
$$\zeta_q^B\left(z^A\right) = 1 - F_q^B\left(z^{A^*} \middle| z^A\right)$$

We let φ^p be the share of players in population p drawing from belief distribution F_1^p, and write μ_q^p for the fraction of players choosing strategy 1 in sub-population q of population p. Letting $\mu \in [0, 1]^4$ be the vector of strategy shares in the sub-populations with typical element μ_q^p, $\varphi \in [0, 1]^2$ the vector of sub-population shares φ^p, and $z(\mu, \varphi) \in [0, 1]^2$ the vector of population states with typical element $z^p(\mu, \varphi) \equiv \varphi^p \mu_1^p + (1 - \varphi^p)\mu_2^p$, the best-response node can be written as

$$\epsilon \begin{pmatrix} \dot{\mu}_1^A \\ \dot{\mu}_2^A \\ \dot{\mu}_1^B \\ \dot{\mu}_2^B \end{pmatrix} = \lambda \cdot \begin{pmatrix} \zeta_1^A(z^B(\mu, \varphi)) - \mu_1^A \\ \zeta_2^A(z^B(\mu, \varphi)) - \mu_1^A \\ \zeta_1^B(z^A(\mu, \varphi)) - \mu_1^B \\ \zeta_2^B(z^A(\mu, \varphi)) - \mu_2^B \end{pmatrix} \tag{11}$$

As in the last section, let average sub-population utility in sub-population q of population p be $\overline{U}_q^p(\mu,\varphi) \equiv \mu_q^p U_1^p(z^{-p}(\mu,\varphi)) + (1-\mu_q^p) U_2^p(z^{-p}(\mu,\varphi))$. Hence, we write the evolution node as

$$\begin{pmatrix} \dot{\varphi}_A \\ \dot{\varphi}_B \end{pmatrix} = \begin{pmatrix} \varphi^A \left(1-\varphi^A\right)\left(\overline{U}_1^A(\mu,\varphi) - \overline{U}_2^A(\mu,\varphi)\right) \\ \varphi^B \left(1-\varphi^B\right)\left(\overline{U}_1^B(\mu,\varphi) - \overline{U}_2^B(\mu,\varphi)\right) \end{pmatrix} \tag{12}$$

As before, we combine the two nodes (11) and (12) to the following slow-fast system:

$$\begin{array}{rcl} \epsilon\dot{\mu} & = & g(\mu,\varphi), \quad \mu \in [0,1]^4 \\ \dot{\varphi} & = & h(\mu,\varphi), \quad \varphi \in [0,1]^2 \end{array} \tag{13}$$

5.2 Stability Condition

Let

$$\mu^*(\varphi) \in \left\{\mu \in (0,1)^4 : \mu = \zeta\left(z(\mu,\varphi)\right)\right\}$$

be a root of $g(\mu,\varphi)$ with typical element $\mu_q^{p*}(\varphi)$, and let $C_{\mu^*}(\varphi) \equiv \partial_\mu g(\mu^*(\varphi),\varphi)$ be the corresponding Jacobi of the best response node. Let now $B_\delta(x) \equiv \{y \in [0,1]^2 : \|x-y\| < \delta\}$. We again focus on belief-monomorphic steady states with $\tilde{\varphi} = (1,1)$. It holds:

Lemma 4. *Fix profile $\{F_1^A, F_2^A, F_1^B, F_2^B\}$ and suppose Assumption 3 (Regularity) holds. Then there is a unique root $\mu^*(\tilde{\varphi})$, and there exists $\delta > 0$ such that*

- *(a) $\mu^*(\tilde{\varphi})$ is differentiable on $B_\delta(\tilde{\varphi})$,*
- *(b) the eigenvalues of $C_{\mu^*}(\varphi')$ all have real parts that are bounded away from zero $\forall \varphi' \in B_\delta(\tilde{\varphi})$, and*
- *(c) the elements $(\mu_1^{A*}(\tilde{\varphi}), \mu_1^{B*}(\tilde{\varphi}))$ of the root $\mu^*(\tilde{\varphi})$ and $C_{\mu^*}(\varphi)$ depend on $\{F_1^A, F_1^B\}$ in $\{F_1^A, F_2^A, F_1^B, F_2^B\}$ alone.*

The proof is left to Appendix A.5. With Lemma 4, the following reduced system is well defined on $B_\delta(\tilde{\varphi})$:

$$\dot{\overline{\varphi}} = h(\mu^*(\overline{\varphi}),\overline{\varphi}), \quad \overline{\varphi} \in B_\delta(\tilde{\varphi})$$

As in the base model, we consider the effect of small a perturbation (μ_0,φ_0) of the rest point $(\tilde{\mu},\tilde{\varphi}) \in [0,1]^6$, and then look at the resulting trajectories (μ_t,φ_t) of the slow-fast system (13). Noting that Theorem 1 does not depend on the particular dimension of (μ,φ), we invoke it again to characterize the trajectory (μ_t, φ_t) by restricting attention to initial conditions (μ_0, φ_0) where $\mu_0 \in [0,1]^4$ lies sufficiently close $\mu^*(\varphi_0)$. That is, again, we look at perturbations characterized by φ_0 alone, and accordingly adapt the stability criterion from the one-population case to the profile $\{F_1^A, F_1^B\}$ of belief distributions. Let $\tilde{z} = z(\tilde{\mu}, \tilde{\varphi})$ be the unique population state in the belief-monomorphic rest point. By Lemma 4 the rest-point $\tilde{z}$ solely depends on $\{F_1^A, F_1^B\}$. Then:

Definition 6 (Replicator Dynamics Stable Belief Distribution)**.** We say that a profile of belief distributions $\{F_1^A, F_1^B\}$ is replicator dynamics stable at $\tilde{z}$, if for all $\epsilon > 0$ small enough it holds that

(a) for every $\Sigma > 0$, there exists a $\Delta(\Sigma) > 0$ such that if $\|\varphi_0 - \tilde{\varphi}\| \leq \Delta(\Sigma)$, then $\|\varphi_t - \tilde{\varphi}\| \leq \Sigma + O(\epsilon)$ for all $t \geq 0$, and
(b) there exists a $\Delta > 0$ such that if $\|\varphi_0 - \tilde{\varphi}\| \leq \Delta$, then $\lim_{t\to\infty} \varphi_t \in B_{O(\epsilon)}(\tilde{\varphi})$

for any profile of belief distribution $\{F_2^A, F_2^B\} \in \mathscr{F} \times \mathscr{F}$ with both F_2^q being $\tilde{z}$-different from F_1^q.

Theorem 2 carries over to this setup, too, because its proof does not depend on the particular dimension of (μ, φ). As the rest point of the best response dynamics is unique, and the corresponding Jacobi $C_{\mu^*}(\varphi)$ depends on $\{F_1^A, F_1^B\}$ alone at $\tilde{\varphi}$ and has strictly negative real eigenvalues in a neighborhood of $\tilde{\varphi}$ by Lemma 4, we have the following corollary:

Corollary 2. *Consider generalized Matching Pennies, and let $\mathscr{F}$ comply with Assumptions 1 (Differentiability), 2 (Richness) and 3 (Regularity). Then any profile of belief distribution $\{F_1^A, F_1^B\}$ that yields $\mu^*(\tilde{\varphi})$ with $(\mu_1^{A*}(\tilde{\varphi}), \mu_1^{B*}(\tilde{\varphi})) = \tilde{z}$ is replicator dynamics stable at $\tilde{z}$ iff $\tilde{\varphi}$ is an asymptotically stable rest point of the reduced system (5) for any belief distribution profile $\{F_2^A, F_2^B\} \in \mathscr{F} \times \mathscr{F}$ with both F_2^q being $\tilde{z}$-different from F_1^q.*

Consequently, we can restrict attention to the reduced system. This yields the following main result of this section:

Proposition 4. *Consider generalized Matching Pennies, and let $\mathscr{F}$ comply with Assumptions 1 (Differentiability), 2 (Richness) and 3 (Regularity). The profile of belief distribution $\{F_1^A, F_1^B\}$ is replicator dynamics stable at $\tilde{z}$ iff $\tilde{z} = z^*$.*

Proposition 4 states that the Nash equilibrium z^* is the only population state at which a belief distribution can be replicator dynamics stable. The proof of Proposition 4 involves several steps and is left to Appendix A.6. The main idea of the proof is to show that a profile $\{F_1^A, F_1^B\}$ of belief distributions is replicator dynamics stable, if and only if for both $p = A, B$ it holds that

$\zeta_1^p(z^{-p^*}) = z^{p^*}$. Thereby, the proof roughly proceeds as follows: Let $z^{p*}(\varphi) = \varphi^p \mu_1^{p*}(\varphi) + (1-\varphi^p)\mu_2^{p*}(\varphi)$. Then by linearizing the reduced system,

$$\dot{\varphi}^A = \varphi^A\left(1-\varphi^A\right)\left(\mu_1^{A*}(\varphi) - \mu_2^{A*}(\varphi)\right)\left(z^{B*}(\varphi)\,a_A - \left(1-z^{B*}(\varphi)\right)b_A\right)$$
$$\dot{\varphi}^B = \varphi^B\left(1-\varphi^B\right)\left(\mu_1^{B*}(\varphi) - \mu_2^{B*}(\varphi)\right)\left(z^{A*}(\varphi)\,a_B - \left(1-z^{A*}(\varphi)\right)b_B\right)$$

it can be shown that any two belief distributions not producing the Nash-equilibrium share can never yield an asymptotically stable distribution-monomorphic rest point on the reduced system. In order to show this, two cases need to be considered: Firstly, we have the case of two belief distributions producing a rest point such that neither distribution-monomorphic population plays the Nash equilibrium share. In this case, the rest-point is hyperbolic, and hence instability is straightforward to establish. The second case has two belief distributions that happen to produce a rest-point such that one but not both populations play the Nash equilibrium share. Such a rest-point is non-hyperbolic with a neutral and an unstable node. This proves the *only if*-part. For the if-part, the linearisation approach fails all together: If we have two belief distributions producing a rest point such that both populations play the Nash equilibrium shares, then the Jacobi matrix of the rest point has only eigenvalues of zero. Consequently, we need to look at higher order terms to establish the result.

An intuitive explanation for the if-part of the Proposition 4 goes as follows: Assume that the system is in its steady state at the mixed Nash equilibrium, and that this state is brought about by distribution-monomorphic populations. Suppose then that in Population A a tiny fraction of agents having a different belief distribution enters. Two scenarios are possible: Firstly, there is a higher fraction of mutant agents drawing below-threshold level beliefs about z^B than there is among the incumbent agents drawing from the stable belief distribution. That is, mutants play strategy 1 more often than incumbents. If so, the overall share of players choosing strategy 1 in Population A increases. Given no intrusion into Population B, the fraction of agents playing strategy 1 increases, too. In such a situation, the superior belief distribution in Population A is the one producing fewer strategy 1-choices. Since belief distributions are continuous in the adversarial strategy share, there exists a small enough share of intruders such that the agents drawing from the initial belief distribution have an advantage after the intrusion and thus crowd the mutants out again. The second scenario has the mutants in Population A draw below-threshold beliefs less often the initially present agents. If so, the overall share of players choosing strategy 1 decreases, and the argument for this scenario unfolds analogously, although with inverse signs. With intrusion on both sides, similar effects are at work to return the system to its steady state.

6 Relation to the Literature

6.1 Evolutionary Foundations of Behavior

While actions are the object of traditional evolutionary game theory (cf. Weibull 1997 for an introduction), preferences have come under scrutiny in what is now known as the indirect evolutionary approach (the seminal paper is Güth & Yaari 1992): Preferences determine actions, actions determine biological fitness, and this feeds back into the evolution of preferences.

In this indirect evolutionary fashion, the idea of having two processes at two different speeds has been employed, for example, by Sandholm (2001*a*), Dekel et al. (2007), or Alger & Weibull (2013), albeit only implicitly. These papers assume that agents' preferences over outcomes evolve slowly while the the adaptation of equilibrium play, given the current preferences, happens instantaneously. Both complete and incomplete information about preferences are considered. In the latter two papers, the stability concept is essentially static. That is, neither the evolutionary process governing the dynamics of preference distributions nor the adoption of equilibrium play is explicitly modeled. In contrast, Sandholm (2001*b*) explicitly models preference evolution in his base model and discusses an extension that explicitely models the co-evolution of beliefs and strategies. Thereby, the evolution of strategy distributions is modeled in a separate node as we do in our model, but it is imposed that the change in the strategy distributions be such that equilibrium play (given the current distribution of preferences) always holds – except possibly after points of discontinuity in the equilibrium distribution where some exogenously defined speed limit is violated. Sandholm (2001*b*) then argues that the case of infinitely fast adaptation can be understood as the limit of this extension with the speed limit going to infinity. The analysis in this paper can be seen as a variant to this argument in the sense that it, too, aims at modeling the strategy adjustment process explicitly but does so neither by imposing equilibrium play at every point nor by bounding the speed of strategy adaptation in absolute terms. Rather, it gives an explicit two-speed model of the way agents arrive at their strategies with evolution running slowly in the background.

6.2 Models of Learning

The idea of deducing an estimate about the population state from the observed play of others as in the sampling procedure is reminiscent of the fictitious play idea: Fictitious play models have agents who collect the whole or just the individual history of play and deduce the current population state from this growing sample by assuming stationarity. It is well known that fictitious play in zero-sum games under suitably chosen initial weights, although converging to the mixed Nash equilibrium in empirical frequencies, leads to cycles in intended play. In their seminal paper, Fudenberg & Kreps (1993) find that for unperturbed games, intended play of the unique mixed Nash equilibrium is stable if and only if the assessment and behavior rules are chosen in a particular manner such that they produce intended equilibrium behavior at empirical equilibrium frequencies. In order to avoid this problem, Fudenberg & Kreps (1993) employ the idea of purifying mixed strategies by means of random payoff shocks: Minimally disturbing payoffs, they find that agents eventually and approximately learn to play the unique mixed strategy equilibrium and that the convergence is global in 2×2-normal form games. Ellison & Fudenberg (2000) extend this analysis. Hofbauer & Sandholm (2002) and Hofbauer & Hopkins (2005) establish global stability of fictitious play in $n \times n$ zero-sum as well as in $n \times n$ partnership games for an arbitrary n by means of Lyapunov functions. By keeping the belief stochastic, this paper follows a different path. Evolutionary pressure on beliefs is presented as a different rationale for assessment rules that produce the unique mixed strategy equilibrium.

Further, the sampling variant of the best-response process in this paper has a loose similarity with the one found in word-of-mouth learning models as, for example, in Banerjee & Fudenberg (2004). Their basic model is as follows: Agents have two choices, one yielding a strictly higher utility than the other. Agents receive a noisy signal of the utility received from their choice and report it to new agents if asked. The new agents then update their information about the state of the world, and act accordingly. Under relatively mild assumptions on the word-of-mouth technology (a sampling rule with sample size $M > 2$), the population of individuals ends up choosing the better option. Similar models with agents using heuristics to learn about the state of the world include the herding papers by Ellison & Fudenberg (1993) and Ellison & Fudenberg (1995). Spiegler (2006*a*) and Spiegler (2006*b*) models consumers of a medical treatment offered by service providers of unknown quality as using word-of-mouth heuristics and analyzes optimal price-setting behavior. A common trait of these models is that the object of learning is stationary. This makes a direct comparison of the results difficult as the assumption of stationarity in the environment stands in sharp contrast to the ideas presented in this paper, where the state of the world might change over time, as it is dependent on the choices made by others.

6.3 Models of Perception

The understanding of perception as a partly random process fits in well with the long tradition of psychological research on the topic. There are two facts that seem to be undisputed: Firstly, human perception is prone to random errors, and secondly, human beings use heuristics when forming beliefs about their environment (cf. Kahneman & Tversky 1996). People using heuristics take less information than there is actually available into account and apply rules-of-thumb to their information subsets. Assuming that the subsets are chosen randomly, this procedure produces beliefs that are stochastic and might be systematically biased. Nevertheless, there is disagreement whether the use of such heuristics is due to limitations in human cognitive computational capacities, and may thus deemed a deficient phenomena (cf. Kahneman & Tversky 1996) or whether they constitute procedures that may even be superior to other procedures which take more information into account. Indeed, it is argued, that heuristics may even produce beliefs that ar more accurate about the environment (cf. Gigerenzer & Brighton 2009). One of the key assumptions of this paper is that heuristics are neither deficient for computability reasons nor optimal in the sense of producing the most accurate beliefs, but rather that they are the product of evolutionary pressure.

In general, the fast growing literature at the boundary between economics and psychology is driven by the underlying premises that individuals might have non-standard beliefs as well as take non-standard actions – whereby the standard refers to the rational choice paradigm (cf. DellaVigna 2009). With the evolutionary perspective taken in this paper, the all-dominant question of why the former kind of such deviations exist in the first place might be addressed: They are stable under evolutionary pressure in interdependent decision problems.

7 Conclusion

This paper is an attempt to analyze the evolutionary foundations of human beliefs in strategic interactions: The context is one of best response dynamics in population games with revising agents drawing a stochastic belief about the current distribution of strategies. There is evolutionary pressure between belief distributions that is modeled by the standard replicator dynamics. The model is one of two differing time-scales: The best response dynamics runs considerably, yet not infinitely, faster than the evolutionary dynamics.

We consider distribution-monomorphic population states with all agents in a population drawing from the same belief distribution and derive a criterion for the stability of belief distributions and the corresponding strategy distributions that obtain in such states. We call a profile of belief distributions replicator dynamics stable if there are no belief distributions that can enter in any of the populations beyond possibly a small fraction whose whose size vanishes in the speed of the best response dynamics. Two applications, one with respect to sampling procedures, and one extending the model to asymmetric 2×2-normal form games, have been discussed.

The reading of the model as an attempt to capture the evolutionary foundations of beliefs and perception calls for some more attention. The manifold implications of the connection between payoff structure, utility functions and stable belief distributions should be scrutinized in additional applications. Future research could be directed to games of higher dimensional strategy spaces.

A Proofs

A.1 Proof of Lemma 2

Proof. Throughout the proof we fix $\tilde{\varphi} = 1$. We start with a preliminary observation that we need later on: If the root $\mu^*(\tilde{\varphi})$ is stable then $\zeta_1'(\mu_1^*(\tilde{\varphi})) < 1$. This can be directly inferred from the strictly positive determinant $det(C_{\mu^*}(\tilde{\varphi}))) = -(\zeta_1'(\mu_1^*(\tilde{\varphi}))-1) > 0$, and the strictly negative trace $tr(C_{\mu^*}(\tilde{\varphi}))) = \zeta_1'(\mu_1^*(\tilde{\varphi}))-2 < 0$ at the stable root $\mu^*(\tilde{\varphi})$.

We start by arguing that any root $\mu^*(\varphi) \in \Upsilon^*(\varphi)$ is differentiable at $\tilde{\varphi}$. The idea is to look at an extended function $\overline{g} : \mathbb{R}^3 \to \mathbb{R}^2$ with $\overline{g}(\mu, \varphi) = g(\mu, \varphi)$ on $[0,1]^3$, and root $\overline{\mu}^*(\varphi)$ solving $\overline{g}(\overline{\mu}^*(\varphi), \varphi) = 0$. Define $\overline{C}_{\overline{\mu}^*}(\varphi) = \partial_\mu \overline{g}(\overline{\mu}^*(\varphi), \varphi)$. Then, $\overline{C}_{\overline{\mu}^*}(\varphi) = C_{\mu^*}(\varphi)$ holds by assumption. By the implicit function theorem, as $\overline{C}_{\overline{\mu}^*}(\tilde{\varphi})$ is invertible, there is a function $\overline{\mu}(\varphi)$ that is continuously differentiable at $\tilde{\varphi}$ and satisfies $\overline{g}(\overline{\mu}^*(\varphi), \varphi) = 0$ for all φ in a neighborhood of $\tilde{\varphi}$.

To show that any $\mu^*(\varphi) \in \Upsilon^*(\varphi)$ is differentiable at $\tilde{\varphi}$, we now argue, that for any root $\overline{\mu}^*(\varphi)$ that is defined in a neighborhood of $\tilde{\varphi}$, it holds that $\overline{\mu}^*(\varphi) = \mu^*(\varphi)$ in the part of that neighborhood which is in $[0,1]$. If this holds, then the relevant one-sided derivatives of $\mu^*(\varphi)$ exist at $\tilde{\varphi}$. Hence, we must show that $\exists \overline{\epsilon} > 0$ such that for all $0 < \epsilon < \overline{\epsilon}$, we have $\overline{\mu}^*(\tilde{\varphi} - \epsilon) \in [0,1]^2$. We distinguish two cases

1. Suppose $\overline{\mu}^*(\tilde{\varphi}) \in (0,1)^2$. In this case it is clear, that $\exists \overline{\epsilon} > 0$ such that for any $0 < \epsilon < \overline{\epsilon}$, we have $\overline{\mu}^*(\tilde{\varphi} - \epsilon) \in [0,1]^2$. Hence, we have $\mu^*(\varphi) = \overline{\mu}(\varphi)$ for all φ in the part of neighborhood of $\tilde{\varphi}$ which is in $[0,1]$, and hence differentiability of $\mu^*(\varphi)$ at $\tilde{\varphi}$.

2. Suppose $\overline{\mu}^*(\tilde{\varphi}) \in \partial\left([0,1]^2\right)$. If $\overline{\mu}^*(\tilde{\varphi})$ is in the boundary $\partial\left([0,1]^2\right)$ of $[0,1]^2$, we need to be more careful. Applying the implicit function theorem on $\overline{g}(\overline{\mu}^*(\varphi), \varphi) = 0$ at $\tilde{\varphi}$, and using the fact that $\overline{\mu}^*(\varphi) = \mu^*(\varphi)$

at $\tilde{\varphi}$, we get for $q = 1, 2$

$$\left.\frac{\mathrm{d}\overline{\mu}_q^*(\varphi)}{\mathrm{d}\varphi}\right|_{\varphi=\tilde{\varphi}} = \frac{\zeta_q'(\mu_1^*(\tilde{\varphi}))(\mu_1^*(\tilde{\varphi}) - \mu_2^*(\tilde{\varphi}))}{1 - \zeta_1'(\mu_1^*(\tilde{\varphi}))}$$

Now, observe that it must hold for any $\mu_1^*(\tilde{\varphi}) \in (0,1)$ that

$$\begin{aligned} \zeta_1(1) = 1 &\Rightarrow \zeta_1'(1) \geq 0 \\ \zeta_1(0) = 0 &\Rightarrow \zeta_1'(0) \geq 0 \\ \zeta_2(\mu_1^*(\tilde{\varphi})) = 1 &\Rightarrow \zeta_1'(\mu_1^*(\tilde{\varphi})) = 0 \\ \zeta_2(\mu_1^*(\tilde{\varphi})) = 0 &\Rightarrow \zeta_1'(\mu_1^*(\tilde{\varphi})) = 0 \end{aligned}$$

As we have $\zeta_1'(\mu_1^*(\tilde{\varphi})) < 1$ and both $\mu_1^*(\tilde{\varphi}), \mu_2^*(\tilde{\varphi})) \in [0,1]$, it then follows that $\left.\frac{\mathrm{d}\overline{\mu}_q^*(\varphi)}{\mathrm{d}\varphi}\right|_{\varphi=\tilde{\varphi}} \geq 0$ if $\mu_q^*(\tilde{\varphi}) = 1$, and $\left.\frac{\mathrm{d}\overline{\mu}_q^*(\varphi)}{\mathrm{d}\varphi}\right|_{\varphi=\tilde{\varphi}} \leq 0$ if $\mu_q^*(\tilde{\varphi}) = 0$. Consequently, $\exists \overline{\epsilon} > 0$ such that $\forall 0 < \epsilon < \overline{\epsilon}$ we have $\overline{\mu}_i^*(\tilde{\varphi} - \epsilon) \in [0,1]$, and hence $\overline{\mu}^*(\tilde{\varphi} - \epsilon) \in [0,1]^2$. Hence, we have $\mu^*(\varphi) = \overline{\mu}(\varphi)$ for all φ in the part of neighborhood of $\tilde{\varphi}$ which is in $[0,1]$, and hence differentiability of $\mu^*(\varphi)$ at $\tilde{\varphi}$.

Having established differentiability of $\mu^*(\varphi)$ at $\tilde{\varphi}$, we now first look at claim (b). The trace and the determinant of $C_{\mu^*}(\varphi)$ for $\varphi \in [0,1]$ are given by

$$\begin{aligned} det(C_{\mu^*}(\varphi)) = 1 - \varphi\zeta_1'\left(\varphi\mu_1^*(\varphi) + (1-\varphi)\mu_2^*(\varphi)\right) \\ - (1-\varphi)\zeta_2\left(\varphi\mu_1^*(\varphi) + (1-\varphi)\mu_2^*(\varphi)\right) \\ tr(C_{\mu^*}(\varphi)) = \varphi\zeta_1'\left(\varphi\mu_1^*(\varphi) + (1-\varphi)\mu_2^*(\varphi)\right) \\ + (1-\varphi)\zeta_2'\left(\varphi\mu_1^*(\varphi) + (1-\varphi)\mu_2^*(\varphi)\right) - 2 \end{aligned}$$

Note that from the fact that both ζ_q', $q = 1, 2$ are continuous by Assumption 1, and both $\mu_q^*(\varphi)$, $q = 1, 2$ are continuous at $\tilde{\varphi}$ by the preceding, it follows that both the trace and the determinant are continuous at $\tilde{\varphi}$. Combined with the fact that the trace $tr(C^*(\varphi))$ is strictly negative at $\tilde{\varphi}$, and the determinant $det(C^*(\varphi))$ is strictly positive at $\tilde{\varphi}$, claim (b) then follows.

Now we turn to claim (a). We now look at $\varphi' \in (0,1)$ in a neighborhood of $\tilde{\varphi}$. If $\mu^*(\varphi') \in (0,1)^2$ the claim follows immediately by the implicit function theorem. When $\mu^*(\varphi')$ is in the boundary of $[0,1]^2$, we need again to be more careful. Let $z^*(\varphi) = \varphi\mu_1^*(\varphi) + (1-\varphi)\mu_2^*(\varphi)$. By the same argument as above, we consider the extended version $\overline{g}$ of g, and arrive at

$$\left.\frac{\mathrm{d}\overline{\mu}_q^*(\varphi)}{\mathrm{d}\varphi}\right|_{\varphi=\varphi'} = \frac{\zeta_q'(z^*(\varphi'))(\mu_1^*(\varphi') - \mu_2^*(\varphi'))}{1 - \varphi'\zeta_1'(z^*(\varphi')) - (1-\varphi')\zeta_2'(z^*(\varphi')} \tag{14}$$

where the denominator is equal to $det(C_{\mu^*}(\varphi'))$, and hence positive. Now, observe that it must hold for $z^*(\varphi') \in (0,1)$, $q = 1,2$,

$$\zeta_q(z^*(\varphi')) \in \{0,1\} \Rightarrow \zeta_q'(z^*(\varphi')) = 0$$

It then follows that $\mathrm{d}\overline{\mu}_q^*(\varphi)/\mathrm{d}\varphi\Big|_{\varphi=\varphi'} = 0$ whenever $\mu_q^*(\varphi') \in \{0,1\}$. Consequently, $\exists \overline{\epsilon} > 0$ such that $\forall -\overline{\epsilon} < \epsilon < \overline{\epsilon}$ we have $\overline{\mu}_i^*(\varphi' - \epsilon) \in [0,1]$, and hence $\overline{\mu}^*(\varphi' - \epsilon) \in [0,1]^2$. Hence, we have $\mu^*(\varphi') = \overline{\mu}(\varphi')$ for all φ in the neighborhood of φ', and hence differentiability of $\mu^*(\varphi)$ at φ'. As we can repeat the argument for all φ' in a neighborhood of $\tilde{\varphi}$, the claim then follows. □

A.2 Proof of Proposition 1

Proof. As we assume that $C^*(\tilde{\varphi})$ has eigenvalues that are strictly below zero, we can focus on the reduced system by Theorem 2. We first look at pure strategy profiles $\tilde{z} \in \{0,1\}$. In order that the pure strategy $\tilde{z} = 1$ ($\tilde{z} = 0$) is stable under the replicator dynamics, we must have that $U_1(\tilde{z}) - U_2(\tilde{z}) > (<)\ 0$. On the other hand, in order that a belief distribution F_1 is replicator dynamics stable at $\tilde{z} = 1$ ($\tilde{z} = 0$), it follows from the reduced system (9) that we must have $-(\mu^{A*}(\tilde{\varphi}) - \mu^{B*}(\tilde{\varphi}))[U_1(z^*(\tilde{\varphi})) - U_2(z^*(\tilde{\varphi}))] < 0$. This is equivalent to $[U_1(z^*(\tilde{\varphi})) - U_2(z^*(\tilde{\varphi}))] > (<)\ 0$ because $\mu^{A*}(\tilde{\varphi}) = \tilde{z} = 1$ and $\mu^{B*}(\tilde{\varphi}) < (>)\ \tilde{z}$. As $\mu^{A*}(\tilde{\varphi}) = z^*(\tilde{\varphi})$, the two requirements for stability are equal.

Next, we look at interior equilibria $\tilde{z} \in (0,1)$. We present two auxiliary results from which the claim of the Proposition then follows:

Lemma 5. *Suppose the eigenvalues of $C_{\mu^*}(\tilde{\varphi})$ all have real parts strictly below zero. If $\zeta_1(\mu_1^*(\tilde{\varphi})) > (<)\ \zeta_2(\mu_1^*(\tilde{\varphi}))$ then $\mathrm{d}z^*(\varphi)/\mathrm{d}\varphi > (<)\ 0$ at $\tilde{\varphi}$.*

Proof. Consider $z^*(\varphi) = \varphi\mu_1^*(\varphi) + (1-\varphi)\mu_2^*(\varphi)$. At $\varphi = \tilde{\varphi}$ we have

$$\frac{\mathrm{d}z^*(\varphi)}{\mathrm{d}\varphi}\bigg|_{\varphi=\tilde{\varphi}} = \mu_1^*(\tilde{\varphi}) - \mu_2^*(\tilde{\varphi}) + \frac{\mathrm{d}\mu_1^*(\varphi)}{\mathrm{d}\varphi}\bigg|_{\varphi=\tilde{\varphi}}$$

Applying the implicit function theorem on $g(\mu^*(\varphi),\varphi) = 0$ at $\tilde{\varphi}$, we get

$$\frac{\mathrm{d}\mu_1^*(\varphi)}{\mathrm{d}\varphi}\bigg|_{\varphi=\tilde{\varphi}} = \frac{\zeta_1'(\mu_1^*(\tilde{\varphi}))(\mu_1^*(\tilde{\varphi}) - \mu_2^*(\tilde{\varphi}))}{1 - \zeta_1'(\mu_1^*(\tilde{\varphi}))}$$

Consequently, we have

$$\frac{\mathrm{d}z^*(\varphi)}{\mathrm{d}\varphi}\bigg|_{\varphi=\tilde{\varphi}} = \frac{\mu_1^*(\tilde{\varphi}) - \mu_2^*(\tilde{\varphi})}{1 - \zeta_1'(\mu_1^*(\tilde{\varphi}))} \tag{15}$$

As $\mu_q^*(\tilde{\varphi}) = \zeta_q(\mu_1^*(\tilde{\varphi})$ and the denominator in (15) is positive because (by the observation in the proof to Lemma 2) the eigenvalues of $C_{\mu^*}(\tilde{\varphi})$ all have negative real parts, the claim follows. □

If a belief distribution F_1 is replicator dynamics stable at $\tilde{z} \in (0,1)$, it follows from the reduced system (9) that $\exists \overline{\varphi} < 1$ such that $\forall \varphi \in [\overline{\varphi}, 1)$, it holds that

$$\varphi(1-\varphi)(\mu_1^*(\varphi) - \mu_2^*(\varphi))[U_1(z^*(\varphi)) - U_2(z^*(\varphi))] > 0 \tag{16}$$

The next lemma is needed to complete the proof:

Lemma 6. *Suppose Assumption 2 (Richness) holds. If a belief distribution F_1 is replicator dynamics stable at $\tilde{z} \in (0,1)$ is stable under random beliefs, it holds that $U_1(\tilde{z}) = U_2(\tilde{z})$.*

Proof. Suppose not, that is, suppose that a belief distribution F_1 is replicator dynamics stable at $\tilde{z} \in (0,1)$, but it holds that $U_1(\tilde{z}) > U_2(\tilde{z})$. As $z^*(\varphi)$ is continuous in φ, there exists (by Assumption 2) a $\mu_2^*(\varphi) > \mu_1^*(\varphi)$ such that the inequality in equation (16) is reversed. We have a contradiction. An analogous argument establishes that $U_1(\tilde{z}) < U_2(\tilde{z})$ cannot hold. □

Consequently, as U_i is differentiable, we need to look at the sign of $U_1'(z) - U_2'(z)$. We consider two cases: (1) $\mu_1^*(\tilde{\varphi}) - \mu_2^*(\tilde{\varphi}) < 0$, and (2) $\mu_1^*(\tilde{\varphi}) - \mu_2^*(\tilde{\varphi}) > 0$.

1. If $\mu_1^*(\tilde{\varphi}) - \mu_2^*(\tilde{\varphi}) < 0$, then $z^*(\varphi) > z^*(\tilde{\varphi})$ for all $\varphi \in (\overline{\varphi}, \tilde{\varphi})$ by Lemma 5. In order that (16) holds, it must hence hold that $U_1'(\tilde{z}) - U_2'(\tilde{z}) < 0$.

2. If $\mu_1^*(\tilde{\varphi}) - \mu_2^*(\tilde{\varphi}) > 0$, then $z^*(\varphi) < z^*(\tilde{\varphi})$ for all $\varphi \in (\overline{\varphi}, \tilde{\varphi})$ by Lemma 5. In order that (16) holds, it must hence hold that $U_1'(\tilde{z}) - U_2'(\tilde{z}) < 0$.

Hence, in order that a belief distribution F_1 is replicator dynamics stable at $\tilde{z} \in (0,1)$ it must hold that $U_1'(\tilde{z}) - U_2'(\tilde{z}) < 0$. Note that this is exactly what is required for replicator dynamics stability of an interior strategy distribution $\tilde{z}$. Hence, we have the claim. □

A.3 Proof of Proposition 2

Proof. We want to show that there exists $F_1 \in \mathscr{F}$ such that the following jointly holds:

(i) $\zeta_1(z^*) = z^*$
(ii) $\zeta_1'(z^*) < 1$

Note that these two points ensure replicator dynamics stability of $F_1 \in \mathscr{F}$ at $z^* \in \{0,1\}$: Point (i) ensures that z^* is indeed a rest point of the best response node in a rule-monomorphic population, and (ii) ensures that the eigenvalues of $C(\mu^*(\tilde{\varphi}), \tilde{\varphi})$ all have real parts strictly below zero. Replicator dynamics stability of $F_1 \in \mathscr{F}$ at $z^* \in \{0,1\}$ then follows by Proposition 1 (as z^* is a replicator dynamics stable strategy).

Because the space of belief distributions brought about by random sampling is rich, point (i) is always satisfied. As $z^* \in \{0,1\}$, it is easily seen from (10) that we must have $p(z^*) = z^*$. Turning to point (ii), we write

$$\zeta_1'(z^*) = p'(z^*) \sum_{i=\lceil M\cdot\overline{z}\rceil}^{M} \binom{M}{i} p(z^*)^{i-1}(1-p(z^*))^{M-i-1}[i - Mz^*]$$

where we have, using the fact that $p(z^*) = z^*$,

$$\zeta_1'(0) = p'(z^*) \left[\sum_{i=\lceil M\cdot\overline{z}\rceil}^{M} \binom{M}{i} i 0^{i-1} \right] = \begin{cases} 0 & \text{if } \lceil M\cdot\overline{z}\rceil > 1 \\ p'(z^*)M & \text{if } \lceil M\cdot\overline{z}\rceil \le 1 \end{cases}$$

and

$$\begin{aligned} \zeta_1'(1) &= p'(z^*) \left[\sum_{i=\lceil M+\cdot\overline{z}\rceil}^{M} \binom{M}{i} [i-M] 0^{M-i-1} \right] \\ &= \begin{cases} 0 & \text{if } \lceil M\cdot\overline{z}\rceil \le M-1 \iff \lceil M\cdot(1-\overline{z})\rceil \ge 1 \\ p'(z^*)M & \text{if } \lceil M\cdot\overline{z}\rceil > M-1 \iff \lceil M\cdot(1-\overline{z})\rceil < 1 \end{cases} \end{aligned}$$

Hence, $\zeta_1'(z^*)$ is either equal to $p'(z^*)\cdot M$, or 0, depending on z^* and $\overline{z}$. Consequently, for any $M \in \mathbb{N}_{++}$, there is a p with $p'(z^*)$ sufficiently small, such that $\zeta_1'(z^*) < 1$ holds. □

A.4 Proof of Proposition 3

Proof. We start with two preliminary observations: (1) For a.e. $\overline{z}$, we have $\lfloor M\overline{z}\rfloor$ invariant, and hence continuous, in $\overline{z}$, and consequently (2) for a.e. $\overline{z}$ we have $I_{1-\overline{z}}(M - \lfloor M\overline{z}\rfloor, \lfloor M\overline{z}\rfloor + 1)$ continuously decreasing in $\overline{z}$. We then need the following auxiliary lemma:

Lemma 7. *Fix some $M \in \mathbb{N}_{++}$, and pick a $\overline{z}$ such that $\overline{z} = I_{1-\overline{z}}(M - \lfloor M\overline{z}\rfloor, \lfloor M\overline{z}\rfloor + 1)$. Then $\exists \delta > 0$ such that $\forall z'$ satisfying $\|z' - \overline{z}\| < \delta$, there is no $M' \le \overline{M}$ such that $z' = I_{1-z'}(M' - \lfloor M'z'\rfloor, \lfloor M'z'\rfloor + 1)$.*

Proof. Take any $\overline{z} \in [0,1]$, and suppose that there exist $M, M' \le \overline{M}$, $M \neq M'$, such that

$$I_{1-\overline{z}}(M - \lfloor M\overline{z} \rfloor, \lfloor M\overline{z} \rfloor + 1) = I_{1-\overline{z}}(M' - \lfloor M'\overline{z} \rfloor, \lfloor M'\overline{z} \rfloor + 1) \tag{17}$$

Since $\partial I_{1-\overline{z}}(M - \lfloor M\overline{z} \rfloor, \lfloor M\overline{z} \rfloor + 1)/\partial \overline{z} \neq \partial I_{1-\overline{z}}(M' - \lfloor M'\overline{z} \rfloor, \lfloor M'\overline{z} \rfloor + 1)/\partial \overline{z}$ for almost every $\overline{z} \in [0,1]$ (the derivative can be shown to exist for almost every $\overline{z}$), it follows that for almost every $\overline{z}$, there exists $\epsilon > 0$ such that for every $z' \in B_\epsilon(\overline{z})$, it holds that

$$I_{1-z'}(M - \lfloor Mz' \rfloor, \lfloor Mz' \rfloor + 1) \neq I_{1-z'}(M' - \lfloor M'z' \rfloor, \lfloor M'z' \rfloor + 1)$$

This implies that for every pair $M, M' \le \overline{M}$, $M \neq M'$, there exist at most finitely many $\overline{z} \in [0,1]$ such that (17) holds. Furthermore, as we only have finitely many pairs $M, M' \le \overline{M}$, $M \neq M'$, there are only at most finitely many $\overline{z}$ such that there exist $M, M' \le \overline{M}$, $M \neq M'$ such that (17) holds. Consequently,

$$\min_{M' \neq M} |I_{1-\overline{z}}(M - \lfloor M\overline{z} \rfloor, \lfloor M\overline{z} \rfloor + 1) - I_{1-\overline{z}}(M' - \lfloor M'\overline{z} \rfloor, \lfloor M'\overline{z} \rfloor + 1)| > 0$$

holds for almost every $\overline{z}$. Together with the continuity of $I_{1-\overline{z}}(M' - \lfloor M'\overline{z} \rfloor, \lfloor M'\overline{z} \rfloor + 1)$ in $\overline{z}$, the claim then follows. □

Consequently, the set of population states $\overline{z}$ at which belief distributions brought about by locally unbiased sampling rules can be replicator dynamics stable consists entirely of singletons. □

A.5 Proof of Lemma 4

Proof. Under Assumption 3, we have $\mathrm{d}\zeta_q^A(z^B)/\mathrm{d}z^B < 0$, $\forall z^B \in [0,1]$ as well as $\mathrm{d}\zeta_q^B(z^A)/\mathrm{d}z^A > 0$, $\forall z^A \in [0,1]$. From this, it follows that there exists a unique root $\tilde{\mu} = \mu^*(\tilde{\varphi})$ that is interior, i.e. in $(0,1)^4$. Given that all $\zeta_q^p(.)$ are differentiable and that the rest-point $\tilde{\mu}$ is interior, differentiability follows from an analogous argument as in the proof to Lemma 2.

We first look at claim (b). As described in the text, the best-response node is given by $\dot{\mu}_1 = g(\mu, \varphi)$ where we have

$$g(\mu,\varphi) = \begin{pmatrix} \zeta_1^A(\varphi^B \mu_1^B + (1-\varphi^B)\mu_2^B) - \mu_1^A \\ \zeta_2^A(\varphi^B \mu_1^B + (1-\varphi^B)\mu_2^B) - \mu_1^A \\ \zeta_1^B(\varphi^A \mu_1^A + (1-\varphi^A)\mu_2^A) - \mu_1^B \\ \zeta_2^B(\varphi^A \mu_1^A + (1-\varphi^A)\mu_2^A) - \mu_2^B \end{pmatrix} \tag{18}$$

Writing out the Jacobi matrix $\partial_\mu g(\mu, \varphi)$, we see that it is of the form:

$$M = \begin{bmatrix} -1 & 0 & -a \cdot x & -a \cdot (1-x) \\ 0 & -1 & -b \cdot x & -b \cdot (1-x) \\ c \cdot y & c \cdot (1-y) & -1 & 0 \\ d \cdot y & d \cdot (1-y) & 0 & -1 \end{bmatrix}$$

with $a, b, c, d > 0$, and $x, y \in [0, 1]$. To get the eigenvalues λ we write the characteristic polynomial

$$\begin{aligned} det(M - \lambda I) = det &\left(\begin{bmatrix} -1-\lambda & 0 & -a \cdot x & -a \cdot (1-x) \\ 0 & -1-\lambda & -b \cdot x & -b \cdot (1-x) \\ c \cdot y & c \cdot (1-y) & -1-\lambda & 0 \\ d \cdot y & d \cdot (1-y) & 0 & -1-\lambda \end{bmatrix} \right) \\ &= (-1-\lambda)^4 + (c \cdot x + d \cdot (1-x)) \cdot (a \cdot y + b \cdot (1-y))(-1-\lambda)^2 \end{aligned} \tag{19}$$

Substituting $\tilde{\lambda} \equiv -1 - \lambda$, and $\tilde{c} \equiv (c \cdot x + d \cdot (1-x)) \cdot (a \cdot y + b \cdot (1-y)) > 0$, we see that the equation

$$\tilde{\lambda}^4 + \tilde{c} \cdot \tilde{\lambda}^2 = 0$$

has roots $\tilde{\lambda} = \{0, \sqrt{-\tilde{c}}, -\sqrt{-\tilde{c}}\}$. Going back to the eigenvalues $\lambda = -1 - \tilde{\lambda}$, we have that the real parts of the eigenvalues of M are all -1, irrespective of a, b, c, d, x, y, and hence uniformly bounded away from zero in (μ, φ). This proves claim (b).

Furthermore, we see from (18) that the elements $(\mu_1^{A*}(\tilde{\varphi}), \mu_1^{B*}(\tilde{\varphi}))$ of the root $\mu^*(\tilde{\varphi})$ only depend on ζ_1^A and ζ_1^B, and hence on $\{F_1^A, F_1^B\}$ alone. Inspecting the characteristic polynomial (19) additionally reveals that for $x, y = 1$ the determinant only depends on ζ_1^A and ζ_1^B, and hence on $\{F_1^A, F_1^B\}$ alone. As the trace of M is constant, claim (c) follows.

Finally, we turn to (a). Given that the rest-point $\tilde{\mu} = \mu^*(\tilde{\varphi})$ is interior and that $\mu^*(\varphi)$ is differentiable at φ, it follows that $\mu^*(\varphi)$ is interior for all φ in neighborhood of $\tilde{\varphi}$. As all $\zeta_q^p(.)$ are differentiable, differentiability of $\mu^*(\varphi)$ on $\varphi \in (0, 1)$ in a neighborhood of $\tilde{\varphi}$ follows directly from the implicit function theorem, whereas for $\varphi \neq \tilde{\varphi}$ in the boundary of $[0, 1]^2$ in a neighborhood of $\tilde{\varphi}$ it follows from an analogous argument as in the proof to Lemma 2. □

A.6 Proof of Proposition 4

Proof. We need to show that $\tilde{\varphi} = (1, 1)$ is an asymptotically stable rest-point of the reduced system if and only if $\mu^*(\tilde{\varphi})$ corresponds to the mixed Nash equilibrium. To this end, we write out the reduced system as:

$$\dot{\varphi}^A = \varphi^A\left(1-\varphi^A\right)\left(\mu_1^{A*}(\varphi)-\mu_2^{A*}(\varphi)\right)\left(z^{B*}(\varphi)\,a_A-\left(1-z^{B*}(\varphi)\right)b_A\right)$$
$$\dot{\varphi}^B = \varphi^B\left(1-\varphi^B\right)\left(\mu_1^{B*}(\varphi)-\mu_2^{B*}(\varphi)\right)\left(z^{A*}(\varphi)\,a_B-\left(1-z^{A*}(\varphi)\right)b_B\right)$$

Let variables marked with a *tilde* denote equilibrium values. We define $\tilde{\varphi} = (\tilde{\varphi}^A, \tilde{\varphi}^B) \equiv (1,1)$, $\tilde{\mu}_q^p \equiv \mu_q^{p*}(\tilde{\varphi})$, and $\tilde{z}^p \equiv z^{p*}(\tilde{\varphi})$. Then, the Jacobian matrix J^* at the rest point $\tilde{\varphi}$ is given by

$$J^* = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{with} \quad \begin{array}{l} a \equiv -\left(\tilde{\mu}_1^A-\tilde{\mu}_2^A\right)\left(\tilde{z}^B a_A-(1-\tilde{z}^B)b_A\right) \\ b \equiv -\left(\tilde{\mu}_1^B-\tilde{\mu}_2^B\right)\left(\tilde{z}^A a_B-(1-\tilde{z}^A)b_B\right) \end{array}$$

In order to check for stability on the reduced system, we need to distinguish three cases regarding the rest-point value of $\tilde{z}^p$: (i) $\tilde{z}^p \neq z^{p^*}$, $\forall p \in P$, (ii) $\tilde{z}^p = z^{p^*}$ for one $p \in P$ but not the other, and (iii) $\tilde{z}^p = z^{p^*}$, $\forall p \in P$. The nature of the rest point differs among the cases. With $\tilde{z}^p \neq z^{p^*}$, $\forall p \in P$, the argument for stability is straightforward, as the rest point is hyperbolic:

Lemma 8. *Let $\tilde{\varphi} = (1,1)$ and $\tilde{z}^p \neq z^{p^*}$, $\forall p \in P$. Then $\exists\{F_2^A, F_2^B\} \in \mathscr{F} \times \mathscr{F}$ such that $\tilde{\varphi}$ is not an asymptotically stable rest point of the reduced system.*

Proof. The eigenvalues (λ_1, λ_2) of the linearized system around the rest point are given by (a, b). Stability requires that $a, b < 0$. Evidently this is not necessarily given as it depends on the rest point distributions produced by $\{F_1^A, F_2^A, F_1^B, F_2^B\}$: When we fix some profile $\{F_1^A, F_1^B\}$, then by Assumption 2 (Richness), there always exists a profile $\{F_2^A, F_2^B\}$ yielding $\tilde{z}$-different reaction functions such that either or both $a, b > 0$. □

In other words, there is always a sub-population that can intrude – on both sides. That is, given $\tilde{z}_1^p \neq z_1^{p^*}$ and any profile of belief distributions $\{F_1^A, F_1^B\}$ yielding that rest point, we can always find another profile of belief distributions $\{F_2^A, F_2^B\}$ such that the distribution-monomorphic rest point is unstable. Next, a similar argument unfolds for rest points with exactly one population's strategy distribution being equal to the Nash equilibrium distribution:

Lemma 9. *Let $\tilde{\varphi} = (1,1)$ and $\tilde{z}^p = z^{p^*}$, but $\tilde{z}^{-p} \neq z^{-p^*}$. Then $\exists\{F_2^A, F_2^B\} \in \mathscr{F} \times \mathscr{F}$ such that $\tilde{\varphi}$ is not an asymptotically stable rest point of the reduced system.*

Proof. With $\tilde{z}_1^p = z_1^{p^*}$ for p but not for $-p$, we get a zero eigenvalue in J^*: Suppose w.l.o.g that $\tilde{z}_1^B = z_1^{B^*}$. Then, the eigenvalues are given by $\lambda = (0, b)$. If $b > 0$ (as holds by appropriately choosing F_2^B), the system has no asymtotically stable manifold and hence cannot be asymptotically stable. □

Hence for any profile of belief distributions $\{F_1^A, F_1^B\}$ that happens to produce the Nash equilibrium distribution of strategies only on one but not on the other side, we can find a belief distributions $\{F_2^A, F_2^B\}$ such that $\tilde{\varphi}$ is not asymptotically stable on the reduced system. Lemmata 8 and 9 taken together then imply that for any set of belief distributions $\{F_1^A, F_1^B\}$ not leading to the unique Nash equilibrium strategy distribution z^* in both populations, we can find belief distributions $\{F_2^A, F_2^B\}$ such that $\tilde{\varphi}$ is not asymptotically stable on the reduced system. This shows the only if part of the statement.

It remains to show that if $\tilde{z}$ equals the unique Nash equilibrium distribution, we have $\tilde{\varphi}$ as an asymptotically stable rest point. This is done in the next lemma:

Lemma 10. *Let $\tilde{\varphi} = (1,1)$ and $\tilde{z}^p = z^{p^*}$, $\forall p \in P$. Then $\tilde{\varphi}$ is an asymptotically stable rest point of the reduced system for any profile $\{F_2^A, F_2^B\} \in \mathscr{F} \times \mathscr{F}$ where for both $p \in P$, F_2^p is $\tilde{z}$-different from F_1^p.*

Proof. Note that the linearized system becomes non-hyperbolic with both eigenvalues zero. Hence, we will need to look at higher order terms in both dimensions. For the sake of clarity, we shift variables such that the rest point comes to lie at $(0,0)$. Let $u = \varphi^A - \tilde{\varphi}^A$, and $v = \varphi^B - \tilde{\varphi}^B$. Furthermore, we have $\mu_q^p = \tilde{\mu}_q^p + d\mu_q^{p*}(\tilde{\varphi})$ with $d\mu_q^{p*}(\tilde{\varphi}) = u\frac{\partial \mu_q^{p*}(\tilde{\varphi})}{\partial \varphi^A} + v\frac{\partial \mu_q^{p*}(\tilde{\varphi})}{\partial \varphi^B}$. For the total distribution of strategies in population B, we then have $z^p = \tilde{z}^B + dz^B(\tilde{\varphi})$ and it follows at the distribution-monomorphic rest point with $\tilde{\varphi}^p = 1$, that

$$dz^{B*}(\tilde{\varphi}) = v(\tilde{\mu}_1^B - \tilde{\mu}_2^B) + u\frac{\partial \mu_1^{B*}(\tilde{\varphi})}{\partial \varphi^A} + v\frac{\partial \mu_2^{B*}(\tilde{\varphi})}{\partial \varphi^B}$$

Application of the implicit function theorem to $g(\mu^*(\varphi), \varphi) = 0$ as given in (18) yields

$$dz^{B*}(\tilde{\varphi}) = u\frac{\zeta_1^{B'}(\tilde{\mu}_1^{A1})(\tilde{\mu}_1^A - \tilde{\mu}_2^2)}{1 - \zeta_1^{A'}(\tilde{\mu}_1^B)\zeta_1^{B'}(\tilde{\mu}_1^A)} + v\frac{(\tilde{\mu}_1^B - \tilde{\mu}_2^B)}{1 - \zeta_1^{A'}(\tilde{\mu}_1^B)\zeta_1^{B'}(\tilde{\mu}_1^A)}$$

Under our change of variable, Population A's branch of the reduced system can be written as

$$\dot{u} = -u(\tilde{\mu}_1^A - \tilde{\mu}_2^A)dz^{B*}(\tilde{\varphi})(a_A + b_A) + O(u^2, v^2)$$

Combining this with the expression for $dz^{B*}(\tilde{\varphi})$ and computing the other dimension analogously, we arrive at:

$$\dot{u} = -u^2(\tilde{\mu}_1^A - \tilde{\mu}_2^A)^2\zeta_1^{B'}(z^{A^*})\frac{a_A + b_A}{1 - \zeta_1^{A'}(z^{B^*})\zeta_1^{B'}(z^{A^*})}$$

$$
\begin{aligned}
&\quad - uv(\tilde{\mu}_1^A - \tilde{\mu}_2^A)(\tilde{\mu}_1^B - \tilde{\mu}_2^B)\frac{a_A + b_A}{1 - \zeta_1^{A'}(z^{B*})\zeta_1^{B'}(z^{A*})} + O(u^3, v^3) \\
\dot{v} &= -v^2(\tilde{\mu}_1^B - \tilde{\mu}_2^B)^2\zeta_1^{A'}(z^{A*})\frac{a_B + b_B}{1 - \zeta_1^{A'}(z^{B*})\zeta_1^{B'}(z^{A*})} \\
&\quad - uv(\tilde{\mu}_1^A - \tilde{\mu}_2^A)(\tilde{\mu}_1^B - \tilde{\mu}_2^B)\frac{a_B + b_B}{1 - \zeta_1^{A'}(z^{B*})\zeta_1^{B'}(z^{A*})} + O(u^3, v^3)
\end{aligned}
$$

In order to show that this system is asymptotically stable on $(u, v) = \{(u, v) : u, v \leq 0\}$ around $(u, v) = (0, 0)$, note that the system is of the form:

$$
\begin{aligned}
\dot{u} &= au^2 + buv \\
\dot{v} &= cv^2 + duv
\end{aligned}
$$

with $a, c > 0$ and $\text{sgn}(b) \neq \text{sgn}(d)$. Without loss of generality, let $d < 0$. Then $b > 0$ and consequently for any $u_0, v_0 < 0$, we have $\dot{u} > 0$, that is, u returns to the origin. So, we need to take a closer look at the dynamics of v: $\dot{v} > 0$ requires $cv^2 > -duv \Leftrightarrow cv < -du$. Evidently, $\forall u\ \exists v\ s.t.\ cv < -du$. Let $v(u) \equiv \sup\{v : cv < -du, u < 0, v < 0\}$ and observe that $\frac{dv(u)}{du} > 0$ holds. As u returns to the origin from below, the trajectory of v, $v(t, v_0)$ starting at v_0, is bounded below for any u_0. Call this bound $\delta = \delta(u_0)$. Hence for any $\delta < 0$ there is an $\epsilon < 0$ such that $v_0 \geq \epsilon$ implies $v(t, v_0) \geq \delta$, $\forall t > 0$ as well as $\lim_{t\to\infty} v(t, v_0) = 0$. Together with the fact that $\dot{u} > 0$ for any $u_0, v_0 < 0$, we have asymptotic stability. In other words, for any $u_0 < 0$, $\exists T > 0$ such that $u = u(t, u_0)$, $\forall t > T$, is sufficiently close to zero such that $\dot{v} > 0$; hence, as soon as u becomes sufficiently close to zero again, both $\dot{u}, \dot{v} > 0$ and the system returns to its steady state at $(0, 0)$. The same reasoning applies for $d > 0$. □

We conclude that if the profile of belief distributions $\{F_1^A, F_1^B\}$ yields the unique mixed strategy distribution for $\tilde{z}$ then $\tilde{\varphi}$ is stable on the reduced system. Combining this with Lemmata 8 and 9 we have established that $\tilde{\varphi}$ is an asymptotically stable rest point of the reduced system if and only if $\tilde{z}$ corresponds to the Nash equilibrium z^*. This completes the proof. □

References

Alger, I. & Weibull, J. (2013), 'Homo Moralis-Preference Evolution under Incomplete Information and Assortative Matching', *Econometrica* **81**(6), 2269–2302.

Banerjee, A. & Fudenberg, D. (2004), 'Word-of-Mouth Learning', *Games and Economic Behavior* **46**(1), 1 – 22.

Berglund, N. & Gentz, B. (2006), *Noise-Induced Phenomena in Slow-Fast Dynamical Systems*, Probability and Its Applications, Springer Berlin Heidelberg.

Dekel, E., Ely, J. C. & Yilankaya, O. (2007), 'Evolution of Preferences', *Review of Economic Studies* **74**(3), 685–704.

DellaVigna, S. (2009), 'Psychology and Economics: Evidence from the Field', *Journal of Economic Literature* **47**(2), 315–372.

Ellison, G. & Fudenberg, D. (1993), 'Rules of Thumb for Social Learning', *The Journal of Political Economy* **101**(4), 612–643.

Ellison, G. & Fudenberg, D. (1995), 'Word-of-Mouth Communication and Social Learning', *The Quarterly Journal of Economics* **110**(1), 93.

Ellison, G. & Fudenberg, D. (2000), 'Learning Purified Mixed Equilibria', *Journal of Economic Theory* **90**(1), 84–115.

Fudenberg, D. & Kreps, D. M. (1993), 'Learning Mixed Equilbria', *Games and Economic Behavior* **5**, 320–367.

Gigerenzer, G. & Brighton, H. (2009), 'Homo Heuristicus: Why Biased Minds Make Better Inferences', *Topics in Cognitive Science* (1), 107–143.

Gilboa, I. & Matsui, A. (1991), 'Social Stability and Equilibrium', *Econometrica* **59**(3), pp. 859–867.

Güth, W. & Yaari, M. (1992), An Evolutionary Approach to Explaining Reciprocal Behavior in a Simple Strategic Game, *in* U. Witt, ed., 'Explaining Process and Change – Approaches to Evolutionary Economics', University of Michigan Press, pp. 23–34.

Hofbauer, J. & Hopkins, E. (2005), 'Learning in Perturbed Asymmetric Games', *Games and Economic Behavior* **52**, 133–152.

Hofbauer, J. & Sandholm, W. H. (2002), 'On the Global Convergence of Stochastic Fictitious Play', *Econometrica* **70**(6).

Kahneman, D. & Tversky, A. (1996), 'On the Reality of Cognitive Illusions', *Psychological Review* **103**(3), 582–591.

Kokotovic, P. V. (1984), 'Applications of Singular Perturbation Techniques to Control Problems', *SIAM review* **26**(4), 501–550.

Olver, F. W., Lozier, D. W., Boisvert, R. F. & Clark, C. W. (2010), *NIST Handbook of Mathematical Functions*, 1st edn, Cambridge University Press, New York, NY, USA.

Oyama, D., Sandholm, W. H. & Tercieux, O. (forthcoming), Sampling Best Resonse Dynamics and Deterministic Equilibrium Selection, *Theoretical Economics*.

Ritzberger, K. & Weibull, J. W. (1995), 'Evolutionary Selection in Normal-Form Games', *Econometrica* **63**(6), 1371–1399.

Sandholm, W. H. (2001*a*), 'Almost Global Convergence to p-dominant Equilibrium', *International Journal of Game Theory* **30**, 107–116.

Sandholm, W. H. (2001*b*), 'Preference Evolution, Two-Speed Dynamics, and Rapid Social Change', *Review of Economic Dynamics* **4**(3), 637 – 679.

Spiegler, R. (2006*a*), 'Competition over Agents with Boundedly Rational Expectations', *Theoretical Economics* **1**, 207–231.

Spiegler, R. (2006*b*), 'The Market for Quacks', *The Review of Economic Studies* **73**(4), 1113–1131.

Taylor, P. D. & Jonker, L. B. (1978), 'Evolutionary Stable Strategies and Game Dynamics', *Mathematical Biosciences* **40**, 145 – 156.

Weibull, J. W. (1997), *Evolutionary Game Theory*, MIT Press.