# Tempered Best Response Dynamics

by

Dai Zusai

# Tempered Best Response Dynamics

Dai Zusai[*]

January 28, 2013

**Abstract**

We propose a new deterministic evolutionary dynamic—the tempered best response dynamic (tBRD)—to capture two features of economic decision making: optimization and continuous sensitivity to incentives. That is, in the tBRD, an agent is more likely to revise his action when his current payoff is further from the optimal payoff, and he always switches to an optimal action when revising. The tBRD is a payoff monotone selection like the replicator dynamic, which makes medium and long-run outcomes more consistent with predictions from equilibrium refinement than the BRD in some situations. The technical contribution of the tBRD is continuous sensitivity, which allows us to apply results of a system of piecewise differential equations in order to obtain conditions for uniqueness and stability of solutions.

*Keywords:* best response dynamic, payoff monotonicity, status-quo bias, switching costs, proper equilibrium, piecewise differential equations

*JEL classification:* C62, C73, D03

---
[*]Department of Economics, Temple University, 1301 Cecil B. Moore Ave., RA 873 (004-04), Philadelphia, PA 19122, U.S.A. Tel:+1-215-204-1762. E-mail: `zusai@temple.edu`.

# 1 Introduction

Among various evolutionary dynamics, the *best response dynamic* (BRD) is thought of as a standard dynamic based on optimization (Gilboa and Matsui, 1991; Matsui, 1992; Hofbauer, 1995): whenever an agent gets an opportunity to revise his action, he chooses the action that is currently optimal. An implicit assumption imposed on the BRD is that revision opportunities come at some constant frequency. At each such opportunity, the revising agent must switch to an optimal action, no matter how little the switch improves the agent's payoff. By the same token, large payoff improvements do not lead an agent to switch his action more quickly.

In reality, the smaller the potential gains, the less likely people are to switch their actions. People may not recognize small gains and may miss chances to take advantage of them; for example, when a commuter drives from his home to his workplace, he may not switch from his regular route for a small difference of a few minutes in drive time. These behaviors can be explained by status-quo biases or switching costs. Experimental and empirical research has reported the existence of status-quo bias in real economic choices: see Samuelson and Zeckhauser (1988); Hartman, Doane, and Woo (1991); Madrian and Shea (2001).[1] By fitting reinforcement learning models with experimental data, Erev and Roth (1998) found that persistence of previously taken actions increases the models' prediction power.

This behavioral evidence motivates us to build a new evolutionary dynamic that combines an optimization-based decision rule with payoff sensitivity in the decision to revise. We define the *tempered best response dynamic* (tBRD) by making the frequency of an agent's revision opportunities depend on the payoff difference between the agent's current action and his optimal action, which we call the the *payoff deficit* of the current action. This dynamic can be interpreted as the standard BRD with stochastic status-quo biases: a larger payoff deficit implies a higher probability that the payoff improvement exceeds the status-quo bias and hence a higher probability that the agent switches to the optimal action.

A status-quo bias seems natural for myopic optimizers. In a class of myopic optimization-based dynamics, such as the standard BRD, myopic agents have the 'stationary' belief that the current state continues into the future. Such agents take this belief as a convenience because of their bounded ability to predict the future. For this reason, they might not let their actions fluctuate with small changes of the current state until they become convinced of those changes, i.e. they would postpone changing the action until the state goes so far from the present one that they are convinced that the present action is undesirable. Thus, a myopic player may want to introduce a buffer that inhibits changes of actions. This role is served by a status-quo bias. In addition, a status-quo bias of a person can vary with the surrounding situation. Stochastic status-quo bias would be a good, tractable way to incorporate effects of the non-modeled surrounding environment on recurrent decision making.

The payoff sensitivity of the tBRD makes its solution trajectories differ from those of the standard BRD in significant ways. Under the BRD, since the switching rate from a suboptimal action does not depend on the payoff, any suboptimal action loses its players at the same constant rate. But under the tBRD, switching rates from suboptimal actions depend on payoff deficits. A worse action has the number of its players decrease at a greater rate than a better one. Hence, the tBRD is a payoff monotone selection like the replicator dynamic. While it admits multiple solution paths like the BRD, its payoff monotonicity eliminates pathological multiplicity and cycles.

The idea of payoff monotonicity appears in equilibrium refinement as well; a proper equilibrium is the limit of perturbations that are consistent with payoff ordering. We present situations where the limit of an interior convergence path under the tBRD is a proper equilibrium. Compared with the standard BRD, where the limit is trembling-hand perfect,

---

[1] Recently decision theorists have incorporated status-quo biases into the axiomatic framework of choice theory: see Masatlioglu and Ok (2005); Sagi (2006); Ortoleva (2010). In addition, the theory of industrial organization notes the significance of consumers' switching costs in market competition (Klemperer, 1995).

the tBRD narrows down plausible outcomes in such situations. Recall that properness implies sequential and subgame-perfect equilibrium in the corresponding extensive form game. We show an example where the BRD allows convergence to an implausible outcome but the tBRD does not.

At the same time, the tBRD preserves the convergence properties of the BRD. We verify global stability of Nash equilibrium in potential games, where the payoff function is derived as the derivative of a single real-valued function. We also establish global stability of Nash equilibrium in stable games, a generalization of potential games with concave potential functions. This leads us to establish local stability of a regular ESS, which is a state where the payoff function locally satisfies the definition of a stable game. Analogous results are established for the 'canonical' deterministic dynamics including the BRD and the replicator dynamic in Sandholm (2001), Hofbauer and Sandholm (2009), and Sandholm (2010a). So the tBRD has as good convergence properties as any of the standard dynamics from the literature.

One technical contribution of the tBRD is to allow us to analyze an optimization-based dynamic by applying results from the literature on systems of piecewise differential equations. This is possible because the speed of transition in the tBRD continuously decreases as the state approaches an equilibrium. We obtain sufficient conditions for uniqueness of a solution path and local stability of an isolated Nash equilibrium. We can use a linearization technique to investigate local dynamic properties under the tBRD, unlike the BRD.

In other contexts, a number of other authors have used status-quo bias or switching costs to model insensitivity to negligible payoff improvements: Kuzmics (2011) and Norman (2009, 2010) in stochastic evolution with noise; Lipman and Wang (2000, 2009) in repeated games; Lou, Yin, and Lawphongpanich (2010) and Szeto and Lo (2006) in transportation models.[2]

These papers, except Kuzmics (2011), assume a deterministic status-quo bias: an agent maintains a threshold over time, and he revises his action if the payoff improvement exceeds this threshold. With a deterministic status-quo bias, an agent is still discontinuously sensitive to payoff differences; he switches from a suboptimal action suddenly, whenever its payoff deficit exceeds some certain threshold. In this respect, these models resemble the standard BRD more than the tBRD, which is derived from a stochastic status-quo bias. Interestingly, while a deterministic status-quo bias can create non-Nash rest points, under a stochastic status-quo bias the rest points of the dynamic are exactly the Nash equilibria of the game being played.

This paper proceeds as follows. In the next section, we introduce population games and review the BRD for reference. In Section 3, we build the tBRD from an exogenous revision rate function and interpret it as a BRD with a stochastic status-quo bias and as a version of fictitious play. In Section 4, we verify stationarity of Nash equilibrium and its stability in various classes of games. In Section 5, we investigate implications of payoff monotonicity of the tBRD with an emphasis on contrasts with the BRD. In Section 6, by taking the tBRD as a system of piecewise differential equations, we establish uniqueness and stability conditions. In the last section, we discuss general implications and possible applications of the tBRD.

## 2   Population games and the best response dynamic

### 2.1   Population games

We consider finite-action games played in a society composed of $P$ populations $\mathcal{P} :=\{1, \ldots, P\}$ of anonymous agents.[3] Each population is a unit mass of agents with the same

---

[2]Among deterministic learning dynamics for mixed strategies with finitely many players, we could view the target projection dynamic (Tsakas and Voorneveld, 2009) as a mixed-strategy (monomorphic) BRD with a deterministic status-quo bias.

[3]For further details, see Sandholm (2010b, Ch.2).

action set and the same payoff function.[4] Each agent in population $p \in \mathcal{P}$ chooses an action $a$ from $\mathcal{A}^p := \{1, \ldots, A^p\}$. Let $A := \sum_{p \in \mathcal{P}} A^p$ equal the total number of actions in all populations.

Denote by $x_a^p \in [0, 1]$ the mass of action-$a$ players in population $p$. The state of population $p$ is represented by a column vector $\mathbf{x}^p := (x_1^p, \ldots, x_{A^p}^p)$ in $\mathcal{X}^p := \Delta\mathcal{A}^p = \{\mathbf{x}^p \in [0, 1]^{A^p} | \mathbf{1} \cdot \mathbf{x}^p = 1\}$.[5] The social state is represented by a column vector $\mathbf{x} := (\mathbf{x}^1, \ldots, \mathbf{x}^P)$ in $\mathcal{X} := \prod_{p \in \mathcal{P}} \mathcal{X}^p$.[6] When we consider only a single population, we omit superscripts for $p = 1$.

The payoff of each action is a function of the social state. Given the state $\mathbf{x} \in \mathcal{X}$, $F_a^p(\mathbf{x})$ is the payoff for a player of action $a \in \mathcal{A}^p$ in population $p$. Define payoff functions $\mathbf{F}^p : \mathcal{X} \to \mathbb{R}^{A^p}$ and $\mathbf{F} : \mathcal{X} \to \mathbb{R}^A$ by

$$\mathbf{F}^p(\mathbf{x}) := \begin{pmatrix} F_1^p(\mathbf{x}) \\ \vdots \\ F_{A^p}^p(\mathbf{x}) \end{pmatrix}, \ \mathbf{F}(\mathbf{x}) := \begin{pmatrix} \mathbf{F}^1(\mathbf{x}) \\ \vdots \\ \mathbf{F}^P(\mathbf{x}) \end{pmatrix} \quad \text{for each } p \in \mathcal{P} \text{ and } \mathbf{x} \in \mathcal{X}.$$

In summary, a population game is mathematically defined by $\mathcal{P}, \mathcal{A}$ and $\mathbf{F}$. Henceforth, we consider games with continuously differentiable payoff functions.

**Assumption 1.** The payoff function $\mathbf{F} : \mathcal{X} \to \mathbb{R}^A$ is continuously differentiable.

Let $b^p(\mathbf{x}) \subset \mathcal{A}^p$ be the set of (pure) best responses for an agent in population $p$ at state $\mathbf{x}$, and let $F_*^p(\mathbf{x}) \in \mathbb{R}$ be the maximized payoff at this state:

$$b^p(\mathbf{x}) := \arg\max_{a \in \mathcal{A}^p} F_a^p(\mathbf{x}), \quad F_*^p(\mathbf{x}) := \max_{a \in \mathcal{A}^p} F_a^p(\mathbf{x}) \quad \text{for each } p \in \mathcal{P}, \mathbf{x} \in \mathcal{X}.$$

The Cartesian product $b(\mathbf{x}) := \prod_{p \in \mathcal{P}} b^p(\mathbf{x}) \subset \mathcal{A}$ collects all possible combinations of optimal actions over all populations. Let $B^p(\mathbf{x}) \subset \Delta\mathcal{A}^p$ and $B(\mathbf{x}) \subset \Delta\mathcal{A}$ be the sets of mixed best responses for population $p \in \mathcal{P}$ and for the whole society, respectively:[7]

$$B^p(\mathbf{x}) := \{\mathbf{y}^p \in \Delta\mathcal{A}^p | y_a^p > 0 \Rightarrow a \in b^p(\mathbf{x})\}, \ B(\mathbf{x}) := \prod_{p \in \mathcal{P}} B^p(\mathbf{x}) \quad \text{for each } p \in \mathcal{P}, \mathbf{x} \in \mathcal{X}.$$

$B^p(\mathbf{x})$ contains all action distributions of population $p$ in which (almost) every agent in $p$ takes actions that are optimal given the payoff vector $\mathbf{F}^p(\mathbf{x})$.

The region in $\mathcal{X}$ where action $a \in \mathcal{A}^p$ is best response of a population-$p$ agent is called the **best response region** of action $a$ and is denoted by $(b^p)^{-1}(a)$. We use similar notation for the best response region of action profile $\mathbf{a} = (a^1, \ldots, a^P) \in \mathcal{A}$:

$$(b^p)^{-1}(a) := \{\mathbf{x} \in \mathcal{X} | a \in b^p(\mathbf{x})\} \quad \text{for each } p \in \mathcal{P}, a \in \mathcal{A}^p,$$

$$b^{-1}(\mathbf{a}) := \{\mathbf{x} \in \mathcal{X} | \mathbf{a} \in b(\mathbf{x})\} = \bigcap_{p \in \mathcal{P}} (b^p)^{-1}(a) \quad \text{for each } \mathbf{a} = (a^1, \ldots, a^P) \in \mathcal{A}.$$

The **payoff deficit** of action $a \in \mathcal{A}^p$ at state $\mathbf{x}$, denoted by $\breve{F}_a^p(\mathbf{x})$, is the difference between the payoffs from action $a$ and from an optimal action:

$$\breve{F}_a^p(\mathbf{x}) := F_*^p(\mathbf{x}) - F_a^p(\mathbf{x}) \quad \text{for each } p \in \mathcal{P}, a \in \mathcal{A}^p, \mathbf{x} \in \mathcal{X}.$$

$\breve{F}_a^p(\mathbf{x})$ is non-negative, and $\breve{F}_a^p(\mathbf{x}) = 0$ means that $a \in b^p(\mathbf{x})$. Let $\bar{F}^p(\mathbf{x})$ be the weighted average of payoffs in population $p$:

$$\bar{F}^p(\mathbf{x}) := \mathbf{F}^p(\mathbf{x}) \cdot \mathbf{x}^p = \sum_{a \in \mathcal{A}^p} x_a^p F_a^p(\mathbf{x}) \quad \text{for each } p \in \mathcal{P} \text{ and } \mathbf{x} \in \mathcal{X}.$$

---

[4]The assumption of unit mass is made just for notational simplicity. We could easily extend the model and the results to general cases where different populations have different masses.

[5]We omit the transpose when we write a column vector on the text. The vector in a bold font is a column vector, while the one with an arrow over the letter is a row vector. $\mathbf{1}$ is a column vector $(1, 1, \ldots, 1)$. Note that $\mathbf{1} \cdot \mathbf{z} = \sum_{i=1}^n z_i$ for an arbitrary column vector $\mathbf{z} = (z_i)_{i=1}^n$. For a finite set $\mathcal{Z} = \{1, \ldots, Z\}$, we define $\Delta\mathcal{Z}$ as $\Delta\mathcal{Z} := \{\rho \in [0, 1]^Z | \mathbf{1} \cdot \rho = 1\}$, i.e. the set of all probability distributions on $\mathcal{Z}$.

[6]Precisely $\mathbf{x}$ is an $A$-dimensional column vector $(x_1^1, \ldots, x_{A^1}^1, x_1^2, \ldots, x_{A^2}^2, \ldots, x_1^P, \ldots, x_{A^P}^P)$.

[7]Notice that $B^p(\mathbf{x}) = \arg\max_{\mathbf{y}^p \in \Delta\mathcal{A}^p} \mathbf{y}^p \cdot \mathbf{F}^p(\mathbf{x})$ is a convex set for every $\mathbf{x} \in \mathcal{X}$.

As usual, a Nash equilibrium is a state where (almost) every agent takes an optimal action. Formally, a social state $\mathbf{x} = (\mathbf{x}^p)_{p \in \mathcal{P}} \in \mathcal{X}$ is a **Nash equilibrium** if the population state $\mathbf{x}^p$ of every population $p \in \mathcal{P}$ is in $B^p(\mathbf{x})$. For each $p \in \mathcal{P}$, the statement $\mathbf{x}^p \in B^p(\mathbf{x})$ is equivalent to $\forall a \in \mathcal{A}^p [x_a^p > 0 \Rightarrow \breve{F}_a^p(\mathbf{x}) = 0]$. Let NE($\mathbf{F}$) be the set of Nash equilibria of $\mathbf{F}$.

The simplest example of a population game is single-population random matching in a symmetric two-player normal-form game with an $n \times n$ payoff matrix $\mathbf{\Pi}$; the population game is defined by $\mathcal{P} = \{1\}, \mathcal{A} = \{1, 2, \ldots, n\}$, and $\mathbf{F}(\mathbf{x}) = \mathbf{\Pi}\mathbf{x}$. NE($\mathbf{F}$) coincides with the set of *symmetric* Nash equilibria of $\mathbf{\Pi}$. Similarly, we can define two-population random matching in a (general) two-player normal-form game. For each $p \in \mathcal{P} = \{1, 2\}$, let $\mathcal{A}^p$ be the set of player $p$'s pure strategies and $U_{ab}^p$ be $p$'s payoff when 1 chooses strategy $a \in \mathcal{A}^1$ and 2 chooses $b \in \mathcal{A}^2$. $U_{ab}^p$ is the $(a, b)$-th cell of the matrix $\mathbf{U}^p$. The payoff function $\mathbf{F}$ is given by $\mathbf{F}^1(\mathbf{x}) = \mathbf{U}^1\mathbf{x}^2$ for population 1 and $\mathbf{F}^2(\mathbf{x}) = (\mathbf{U}^2)^T\mathbf{x}^1$ for population 2.[8] NE($\mathbf{F}$) coincides with the set of all (both symmetric and asymmetric) Nash equilibria of the normal-form game $(\mathbf{U}^1, \mathbf{U}^2)$.

## 2.2 The best response dynamic

In an evolutionary model of behavior in a population game, agents play the game recurrently over time; the social state changes as they switch their actions, and this changes the payoff at each moment in time. An individual player only occasionally revises his action; such a revision opportunity follows a Poisson process. The *revision rate* is the arrival rate of this Poisson process. On a revision opportunity, the revising agent follows a *revision protocol*—e.g. myopic optimization, mutation—to decide on a new action based on the current information on the payoff and the social state.[9]

The best response dynamic is defined by a constant revision rate and myopic optimization. Although the idea naturally emerges as a disequilibrium adjustment process in economic models, it was not until the early 1990s that BRD was formalized and investigated as an evolutionary dynamic in a general game theoretic context. Hofbauer (1995) defines the **best response dynamic** (BRD) as[10]

$$\dot{\mathbf{x}} \in B(\mathbf{x}) - \mathbf{x}. \tag{1}$$

To interpret this dynamic, we imagine the following revision process. An individual agent receives a revision opportunity from a rate-1 Poisson process. At each revision opportunity, he chooses one of the pure best responses to the current social state $\mathbf{x}$. The social state gradually moves from $\mathbf{x}$ toward a new action distribution. As we do not restrict how an agent chooses a new action from multiple pure best responses, the distribution of new actions among all the revising agents can be any convex combination of the pure best responses, i.e. any state in $B(\mathbf{x})$. Thus, the BRD may allow multiple feasible transition vectors from a single social state $\mathbf{x}$ and so it is defined as a differential inclusion (a multi-valued differential equation).

Note that, under the best response dynamic, an agent switches to an optimal action at every revision opportunity, regardless of the payoff improvement from the switch. However, in reality, people might ignore or even miss a chance to make small improvements, possibly due to psychological status-quo bias or physical switching costs. On the other hand, people would not likely miss such a chance if their present action is quite disadvantageous. In short, we expect that the higher the payoff deficit, the more likely a change in actions. This motivates us to modify a BRD in a way to make it sensitive to payoff deficits but still based on optimization.

---

[8]Here $M^T$ is the transpose of a matrix $M$.

[9]See Sandholm (2010b, Ch.4). He defines each of the major evolutionary dynamics by these two components and then induces a differential equation/inclusion from the aggregate (the mean dynamic) of the individual revisions.

[10]Gilboa and Matsui (1991) and Matsui (1992) also formulate the best response dynamic in slightly different forms.

# 3 Tempered best response dynamics

To define the protocol for the tempered best response dynamic (tBRD), we first assume directly that each agent's revision rate depends exogenously on his payoff deficit. We discuss that this model is equivalent to one defined by a stochastic status-quo bias. We interpret the tBRD as an approximation of the fictitious play like the BRD. In the end of this section, we confirm the existence of a solution from any initial state.

## 3.1 Defining the tBRD

Keeping myopic optimization for the revision protocol from the standard BRD, now we assume that, given the current payoff vector $\boldsymbol{\pi}^p$, the revision rate of an action-$a$ player in population $p$ is $Q(\check{\pi}_a^p)$ with a function $Q : \mathbb{R}_+ \to [0,1]$ of his current payoff deficit $\check{\pi}_a^p := \max_{b \in \mathcal{A}^p} \pi_b^p - \pi_a^p$. We further assume that the greater the payoff deficit of the current action, the more likely an agent is to abandon this action:

**Assumption 2.** The function $Q : \mathbb{R}_+ \to [0,1]$ is strictly increasing and continuously differentiable.

In addition, to make a clear distinction from the BRD, we also assume the following.

**Assumption 3.** The function $Q : \mathbb{R}_+ \to [0,1]$ satisfies $Q(0) = 0$ and $Q(q) > 0$ for all $q > 0$.

In words, Assumption 3 says that revision opportunities never arrive as long as the agent is taking an optimal action and otherwise he receives an opportunity with some positive probability. By this *tempering* function $Q$, the frequency of revision is tempered in a way sensitive to the payoff deficit, compared to the standard BRD with the revision rate 1.

Now consider how the social state changes in an infinitesimal period of time $[t, t+dt]$ under tempered revision rates and myopic optimization. As the total mass of action-$a$ players is $x_a^{t,p}$ and their revision rates are $Q(\check{F}_a^p(\mathbf{x}^t))$ in the current social state $\mathbf{x}^t$, the mass $x_a^{t,p}Q(\check{F}_a^p(\mathbf{x}^t))dt$ revises actions from action $a$ in this period. Like with the standard BRD, each of them individually switches to any of the pure best responses to the current social state $\mathbf{x}^t$. The new action distribution among them is represented by $\mathbf{y}_a^p \in B^p(\mathbf{x}^t)$, while the old one is $\mathbf{e}_a^p$ as they were all playing action $a$ before this revision. Their revisions change the population state by $x_a^{t,p}Q(\check{F}_a^p(\mathbf{x}^t))dt(\mathbf{y}_a^p - \mathbf{e}_a^p)$. Aggregating such changes over the mass of each action players in population $p$, we obtain the total transition of the population state $\mathbf{x}^p$ in this period:

$$\mathbf{x}^{t+dt,p} - \mathbf{x}^{t,p} = \sum_{a \in \mathcal{A}^p} x_a^{t,p}Q(\check{F}_a^p(\mathbf{x}^t))dt(\mathbf{y}_a^p - \mathbf{e}_a^p) \quad \text{for each } p \in \mathcal{P}.$$

Dividing both sides by $dt$ and taking the limit of $dt \to 0$, we have[11]

$$\dot{\mathbf{x}}^{t,p} = \sum_{a \in \mathcal{A}^p} x_a^{t,p}Q(\check{F}_a^p(\mathbf{x}^t))(\mathbf{y}_a^p - \mathbf{e}_a^p) \in \sum_{a \in \mathcal{A}^p} x_a^{t,p}Q(\check{F}_a^p(\mathbf{x}^t))(B^p(\mathbf{x}^t) - \mathbf{e}_a^p) \quad \text{for each } p \in \mathcal{P}.$$

Define the tempered best response correspondence $B_Q : \mathcal{X} \rightrightarrows \mathcal{X}$ by $B_Q(\mathbf{x}) := \prod_{p \in \mathcal{P}} B_Q^p(\mathbf{x})$ with $B_Q^p : \mathcal{X} \rightrightarrows \mathcal{X}^p$ given by

$$B_Q^p(\mathbf{x}) := \sum_{a \in \mathcal{A}^p} x_a^p \left\{ Q(\check{F}_a^p(\mathbf{x}))B^p(\mathbf{x}) + (1 - Q(\check{F}_a^p(\mathbf{x})))\mathbf{e}_a^p \right\} \subset \mathcal{X}^p \quad \text{for each } p \in \mathcal{P}.$$

---

[11]Roth and Sandholm (2012) consider finite-population optimization-based evolutionary dynamics, including tBRD, both in discrete and continuous time horizons. They prove that, as the size of a population goes to infinity, both the medium and long run behavior of the dynamic is well approximated by the infinite-population dynamic such as the one presented here.

Then we define the **tempered best response dynamic** (tBRD) as the differential inclusion below induced from a tempering function $Q$:[12]

$$\dot{\mathbf{x}} \in B_Q(\mathbf{x}) - \mathbf{x} =: V_Q(\mathbf{x}) \subset T\mathcal{X}. \tag{2}$$

While we have defined the tBRD from tempered revision rate, we can interpret this as a BRD with stochastic status-quo biases. That is, upon each revision opportunity whose arrival rate is one, an agent draws a status-quo bias $q \geq 0$ from the distribution function $Q$.[13] The agent switches to a pure best response $b \in b^p(\mathbf{x})$ only if the switch makes the payoff improvement greater than the status-quo bias $q$, namely if $\check{F}_a^p(\mathbf{x}) > q$ where $a$ is his action before revision. So, before the status-quo bias is drawn, the probability that an action-$a$ agent makes a switch receiving the revision opportunity is just $Q(\check{F}_a^p(\mathbf{x}))$. Therefore, the proportion of those who actually make switches among all action-$a$ players is only $Q(\check{F}_a^p(\mathbf{x}))$, just the same as in the case of the tempered revision rate. As a result, we obtain the same population dynamic as (2).

The payoff sensitivity of the revision rate makes the direction of the transition different from the standard BRD, as it creates a bias in the selection of revising agents. Provided that the average revision rate

$$\tilde{Q}^p(\mathbf{x}) := \sum_{a \in \mathcal{A}^p} x_a^p Q(\check{F}_a^p(\mathbf{x})) \in [0, 1]$$

is not zero, tBRD (2) simplifies to

$$\dot{\mathbf{x}}^p = \tilde{Q}^p(\mathbf{x})(\tilde{\mathbf{y}}^p(\mathbf{x}) - \tilde{\mathbf{x}}^p(\mathbf{x})) \in V_Q^p(\mathbf{x}) = \tilde{Q}^p(\mathbf{x})(B^p(\mathbf{x}) - \tilde{\mathbf{x}}^p(\mathbf{x})), \tag{3}$$

where

$$\tilde{\mathbf{y}}^p(\mathbf{x}) := \sum_{a \in \mathcal{A}^p} \frac{x_a^p Q(\check{F}_a^p(\mathbf{x}))}{\tilde{Q}^p(\mathbf{x})} \mathbf{y}_a^p(\mathbf{F}^p(\mathbf{x})) \in B^p(\mathbf{x}) \text{ and } \tilde{\mathbf{x}}^p(\mathbf{x}) := \sum_{a \in \mathcal{A}^p} \frac{x_a^p Q(\check{F}_a^p(\mathbf{x}))}{\tilde{Q}^p(\mathbf{x})} \mathbf{e}_a^p \in \Delta \mathcal{A}^p.$$

Here $\tilde{x}_a^p = x_a^p Q(\check{F}_a^p(\mathbf{x}))/\tilde{Q}^p(\mathbf{x})$ is the proportion of revising agents who originally played action $a$. So $\tilde{\mathbf{x}}^p = (x_a^p Q(\check{F}_a^p(\mathbf{x}))/\tilde{Q}^p(\mathbf{x}))_{a \in \mathcal{A}^p} \in \Delta \mathcal{A}^p$ is the distribution of actions among revising agents in the population $p$, and $\tilde{\mathbf{y}}^p \in B^p(\mathbf{x})$ is their action distribution *after* the revision. By assumption 2, the rate at which revision opportunities are received is larger for a larger payoff deficit. This causes a selection bias in the distribution of old actions among revising agents, $\tilde{\mathbf{x}}$. In contrast, under the standard BRD (1), the rate at which revision opportunities are received is independent of agents' actions and payoffs. Thus the old action distribution of revising agents is just the same as the social state $\mathbf{x}$. We illustrate this difference geometrically in the following examples.[14]

*Example* 1. Consider a single population with three actions $\mathcal{A} = \{1, 2, 3\}$. Suppose that the social state is $\mathbf{x} = (1/3, 1/3, 1/3)$ and that payoffs satisfy $F_1(\mathbf{x}) > F_2(\mathbf{x}) > F_3(\mathbf{x})$, so that the best response is action 1: $B(\mathbf{x}) = \{\mathbf{e}_1\}$.

Under the BRD, the social state changes towards $\mathbf{e}_1 = (1, 0, 0)$. As we noted before, the action distribution of revising agents, $\tilde{\mathbf{x}}$, coincides with the social state $\mathbf{x}$.

Switching the action to the best response gives more payoff improvement to an action 3 player than an action 2 player. So tBRD gives revision opportunities more quickly to action 3 players than action 2 players, and gives no opportunities to action 1 players: $Q_3 > Q_2 > Q_1 = 0$. So although the current state is $\mathbf{x} = (1/3, 1/3, 1/3)$, the action distribution $\tilde{\mathbf{x}}$ among revising players satisfies

$$\tilde{x}_3 = \frac{Q_3}{Q_3 + Q_2} > \tilde{x}_2 = \frac{Q_2}{Q_3 + Q_2} > \tilde{x}_1 = 0.$$

---

[12]Here $T\mathcal{X} := \prod_{p \in \mathcal{P}} T\mathcal{X}^p$ and $T\mathcal{X}^p$ is the tangent space of $\mathcal{X}^p \subset \mathbb{R}^{A^p}$, i.e. $T\mathcal{X}^p := \{\mathbf{z}^p \in \mathbb{R}^{A^p} | \mathbf{1} \cdot \mathbf{z}^p = 0\}$.

[13]Assumption 2 indeed enables us to interpret $Q$ as a distribution function.

[14]When $\mathbf{x}$ is fixed or clear from the context, we abbreviate $Q(\check{F}_a^p(\mathbf{x}))$ as $Q_a^p$ and $Q'(\check{F}_a^p(\mathbf{x}))$ as $Q_a^{p\prime}$.
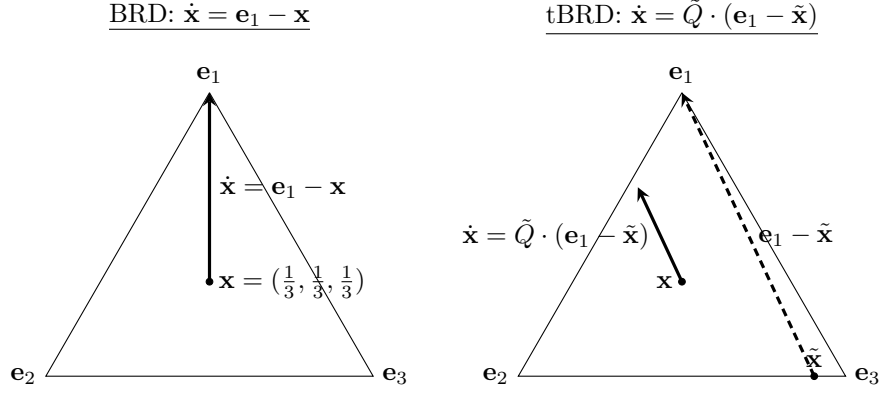
BRD: $\dot{\mathbf{x}} = \mathbf{e}_1 - \mathbf{x}$      tBRD: $\dot{\mathbf{x}} = \tilde{Q} \cdot (\mathbf{e}_1 - \tilde{\mathbf{x}})$

*Figure 1:* Example 1. $F_1(\mathbf{x}) > F_2(\mathbf{x}) > F_3(\mathbf{x})$.



BRD: $\dot{\mathbf{x}} \in B(\mathbf{x}) - \mathbf{x}$      tBRD: $\dot{\mathbf{x}} \in \tilde{Q} \cdot (B(\mathbf{x}) - \mathbf{e}_3)$
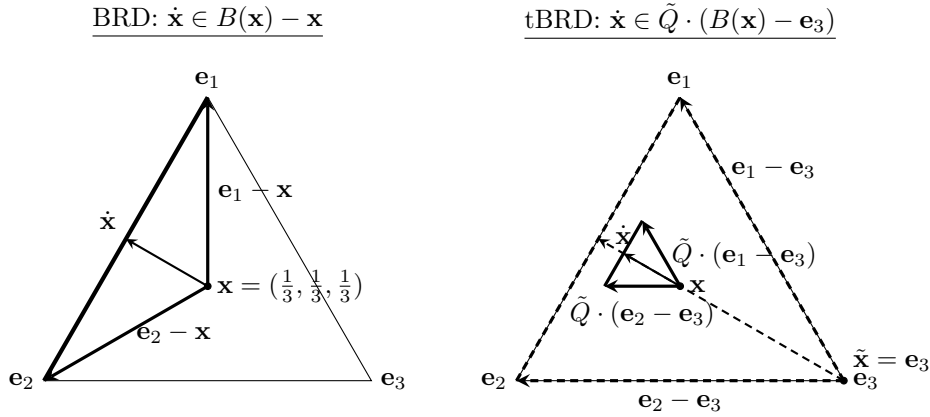
*Figure 2:* Example 2. $F_1(\mathbf{x}) = F_2(\mathbf{x}) > F_3(\mathbf{x})$. Here the mixed best response $B(\mathbf{x})$ is all convex combinations of $\mathbf{e}_1$ and $\mathbf{e}_2$, i.e. $\{\lambda \mathbf{e}_1 + (1-\lambda)\mathbf{e}_2 | \lambda \in [0,1]\}$.

In tBRD, the vector $\mathbf{e}_1 - \tilde{\mathbf{x}}$ is the change in the action distribution over the revising agents. Shrinking this vector by the proportion of revising agents $\tilde{Q}$, we obtain the change in the action distribution of the whole society $\dot{\mathbf{x}} = \tilde{Q} \cdot (\mathbf{e}_1 - \tilde{\mathbf{x}})$. ∎

*Example 2.* Consider the same situation as above, i.e. a single population, $\mathcal{A} = \{1, 2, 3\}$ and $\mathbf{x} = (1/3, 1/3, 1/3)$, but suppose that the payoff ranking is $F_1(\mathbf{x}) = F_2(\mathbf{x}) > F_3(\mathbf{x})$, so that actions 1 and 2 are both best responses.

Under BRD, the social state can move toward any point between $\mathbf{e}_1$ and $\mathbf{e}_2$. Notice that BRD allows action 2 players to revise their actions to action 1 and vice versa. However, such a revision does not improve their payoffs.

With Assumption 3, tBRD does not let action 1 and action 2 players revise their actions. Hence, only action 3 players revise their actions to either action 1 or action 2. So the mass of revising agents consists entirely of action 3 players: $\tilde{\mathbf{x}} = \mathbf{e}_3$. Their new action distribution can be any convex combination between $\mathbf{e}_1$ and $\mathbf{e}_2$. Noticing that the mass of revising agents is just $\tilde{Q} = x_3 Q_3 < 1$, we find that the set of feasible transition vectors $\dot{\mathbf{x}}$ under the tBRD is the set of all convex combinations of $x_3 Q_3 (\mathbf{e}_1 - \mathbf{e}_3)$ and $x_3 Q_3 (\mathbf{e}_2 - \mathbf{e}_3)$. ∎

In general, any two transition vectors from a single social state always form an acute angle under the tBRD, while they may form an obtuse angle under the standard BRD as Figure 2 shows. By (3), any two feasible non-zero transition vectors $\mathbf{v}, \mathbf{w} \in V(\mathbf{x})$ from $\mathbf{x}$ under the tBRD can be written as $\mathbf{v} = \tilde{Q}(\mathbf{x})(\mathbf{y} - \tilde{\mathbf{x}})$ and $\mathbf{w} = \tilde{Q}(\mathbf{x})(\mathbf{z} - \tilde{\mathbf{x}})$ with $\mathbf{y}, \mathbf{z} \in B(\mathbf{x})$.

Let $\theta$ be their angle. Then,

$$\cos\theta = \frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{v}\|\cdot\|\mathbf{w}\|} = \frac{(\mathbf{y}-\tilde{\mathbf{x}})\cdot(\mathbf{z}-\tilde{\mathbf{x}})}{\|\mathbf{y}-\tilde{\mathbf{x}}\|\cdot\|\mathbf{z}-\tilde{\mathbf{x}}\|} = \frac{\mathbf{y}\cdot\mathbf{z} - \mathbf{y}\cdot\tilde{\mathbf{x}} - \mathbf{z}\cdot\tilde{\mathbf{x}} + \|\tilde{\mathbf{x}}\|^2}{\|\mathbf{y}-\tilde{\mathbf{x}}\|\cdot\|\mathbf{z}-\tilde{\mathbf{x}}\|}.$$

By $\tilde{\mathbf{x}}, \mathbf{y}, \mathbf{z} \in \mathcal{X} \subset \mathbb{R}^A_+ \setminus \{\mathbf{0}\}$, we have $\mathbf{y}\cdot\mathbf{z} \geq 0$ and $\|\tilde{\mathbf{x}}\| > 0$. If $y^p_a > 0$, $a$ is an optimal action and $\tilde{x}^p_a = 0$ by Assumption 3; hence, we have $\mathbf{y}\cdot\tilde{\mathbf{x}} = 0$ and similarly $\mathbf{z}\cdot\tilde{\mathbf{x}} = 0$. So we have $\cos\theta > 0$; that is, the angle $\theta$ between the two transition vectors is acute.

## 3.2   Interpretation as fictitious play

It is well known that the standard BRD captures the transition of beliefs in fictitious play learning. So it is natural to interpret the tBRD as a version of fictitious play. In standard fictitious play, agents recursively play a game in a discrete-time setting. In each period, they observe the empirical distribution of past actions; each plays best responses to it, believing that his opponents choose their actions according to this distribution.

For simplicity, consider a single population of continuously many agents who have common belief. Let $\boldsymbol{\theta}^t \in \mathcal{X}$ be the empirical action distribution (belief) at the beginning of period $t$ and $\mathbf{y}^t$ be the distribution of actions taken in period $t$. Then they follow

$$\begin{cases} \boldsymbol{\theta}^t = \dfrac{1}{t-1}\displaystyle\sum_{\tau=1}^{t-1}\mathbf{y}^\tau, \quad \text{i.e. } \boldsymbol{\theta}^t - \boldsymbol{\theta}^{t-1} = \dfrac{1}{t-1}(\mathbf{y}^{t-1} - \boldsymbol{\theta}^{t-1}), \text{ and} \\ \mathbf{y}^t \in B(\boldsymbol{\theta}^t); \end{cases} \tag{4}$$

$$\therefore \quad \boldsymbol{\theta}^{t+1} - \boldsymbol{\theta}^t \in \frac{1}{t}(B(\boldsymbol{\theta}^t) - \boldsymbol{\theta}^t) \text{ for each } t \in \mathbb{N}.$$

Benaïm, Hofbauer, and Sorin (2005) prove that the standard BRD (1) approximates this process of belief under standard fictitious play.

By replacing the best response correspondence $B$ with the tempered one $B_Q$, we obtain

$$\boldsymbol{\theta}^{t+1} - \boldsymbol{\theta}^t \in \frac{1}{t}(B_Q(\boldsymbol{\theta}^t) - \boldsymbol{\theta}^t) \quad \text{for each } t \in \mathbb{N}. \tag{5}$$

Similarly to the standard BRD and fictitious play, the tempered BRD (2) approximates this process.[15]

To interpret (5), we reinterpret fictitious play as follows. At each period, a unit mass of new agents enter the society. They observe past actions made by all the incumbents (not only the last entrants). Then, the entrants play the best response to it. Each agent plays an action only once. At the beginning of period $t$, the total mass of the incumbents is $t-1$ and the proportion of the mass of the entrants among all the agents in the society is just $1/t$. The distribution of the new actions made in period $t$ and the distribution of the past actions observed at the beginning of period $t$ coincide with $\mathbf{y}^t$ and $\boldsymbol{\theta}^t$ in (4).

Now we introduce a status-quo bias to the protocol to decide on a new action. When entering the society, each entrant randomly meets with one of the incumbents and takes her action as a 'default.' He also observes $\boldsymbol{\theta}^t$ and takes it as his belief about his opponents' actions. Then, the entrant draws a status-quo bias from the distribution function $Q$. According to $\boldsymbol{\theta}^t$, if he believes payoff improvement exceeds the status-quo bias, he plays a pure best response; otherwise, he plays the default action. Then $\boldsymbol{\theta}^t$ and $\mathbf{y}^t$ jointly follow

$$\begin{cases} \boldsymbol{\theta}^t = \dfrac{1}{t-1}\displaystyle\sum_{\tau=1}^{t-1}\mathbf{y}^\tau, \quad \text{i.e. } \boldsymbol{\theta}^t - \boldsymbol{\theta}^{t-1} = \dfrac{1}{t-1}(\mathbf{y}^{t-1} - \boldsymbol{\theta}^{t-1}), \text{ and} \\ \mathbf{y}^t \in \displaystyle\sum_{a\in\mathcal{A}}\theta^t_a\left\{Q(\breve{F}_a(\boldsymbol{\theta}^t))B(\boldsymbol{\theta}^t) + (1 - Q(\breve{F}_a(\boldsymbol{\theta}^t)))\mathbf{e}_a\right\}, \end{cases}$$

which reduces to (5).

---

[15]See Benaïm, Hofbauer, and Sorin (2005, Prop. 1.3). Note that here we consider a deterministic process over a discrete time horizon in a population of continuously many agents.

## 3.3 Existence of solution paths

Like with the BRD, existence of solution paths under the tBRD is not guaranteed by standard results for differential equations. This is not only because there may be multiple transition vectors from a single state, but also because the transition vector discontinuously changes when the state crosses the best response regions. So we may not find a solution path that is differentiable at *every* moment in time. To ensure the existence of solutions while allowing change in best responses, we adopt a **Carathéodory solution**. That is, we allow a solution path not to be differentiable and not to have $\dot{\mathbf{x}}^t \in V_Q(\mathbf{x}^t)$ at a measure-zero set of moments in time, while requiring Lipschitz continuity at every time.

Doing so, we can guarantee the existence of a solution path under the tBRD. Note that, unlike a differential equation, uniqueness of a solution is not guaranteed. For example, a typical coordination game, e.g. $\mathbf{F}(\mathbf{x}) = \mathbf{x}$, allows multiple solution paths if there are multiple optimal best responses to the initial state. In Section 6, we present a uniqueness condition that is peculiar to the tBRD.

**Proposition 1.** *Consider a tBRD $\dot{\mathbf{x}} \in V_Q(\mathbf{x})$ and assume Assumptions 1-3. From any state $\mathbf{x} \in \mathcal{X}$, there exists a Carathéodory solution that starts at $\mathbf{x}$ and remains in $\mathcal{X}$ for all positive times.*

*Proof.* The correspondence $V_Q : \mathcal{X} \rightrightarrows T\mathcal{X}$ is nonempty-, closed- and convex-valued, upper semi-continuous and bounded like the best response correspondence. Then, the existence of a solution path is verified by Smirnov (2001, Corollary 4.4). □

# 4 Stationarity and stability of Nash equilibrium

There are several basic properties that a 'reasonable' evolutionary dynamic should satisfy: stationarity of Nash equilibrium and its stability in some classes of games. Here we confirm that the tBRD preserves all these properties. In particular, the tBRD retains the regular sense of Nash stationarity that is lost in the standard BRD. In the last part of this section, we argue that the best response property, shared by the BRD and the tBRD, implies Nash stability distinct from other evolutionary dynamics such as the replicator dynamic.

## 4.1 Nash stationarity and positive correlation

Because of its status-quo bias, tBRD retains the 'standard' Nash stationarity, stronger than that of the BRD. Under the BRD, a stationary 'path' (staying at the same state forever) is always possible at a Nash equilibrium, but the BRD may allow the social state to move away; $\mathbf{x}$ being a Nash equilibrium only implies that $\mathbf{0}$ is one possible transition vector, and not that $\mathbf{0}$ is the only one. Actually, at an interior Nash equilibrium, all actions are optimal; hence, multiple directions of transition are possible.[16] In contrast, the tBRD never admits transition vectors other than $\mathbf{0}$ or any paths other than a stationary path at a Nash equilibrium.

**Theorem 1** (Nash stationarity). *Consider a tBRD $\dot{\mathbf{x}} \in V_Q(\mathbf{x})$; suppose Assumption 3 holds.*

1. *If the state $\mathbf{x} \in \mathcal{X}$ is a Nash equilibrium, then $V_Q(\mathbf{x}) = \{\mathbf{0}\}$.*

2. *If the state $\mathbf{x} \in \mathcal{X}$ is not a Nash equilibrium, then $\mathbf{0} \notin V_Q(\mathbf{x})$.*

3. *In addition, suppose Assumptions 1 and 2 hold. If the state $\mathbf{x}^* \in \mathcal{X}$ is a Nash equilibrium, then the stationary path $\mathbf{x}^t \equiv \mathbf{x}^*$ is the unique Carathèodory solution from $\mathbf{x}^0 = \mathbf{x}^*$.*

---

[16]However, not every direction can be maintained once the state leaves the equilibrium; a transition vector has to be consistent with the best response at off-equilibrium states in the direction of its vector. We can see this in the examples that follow.

*Proof.* See Appendix A.1 □

As well as Nash stationarity, positive correlation is shared by most major evolutionary dynamics and is thought as one of the desiderata that a dynamic should satisfy for consistency with incentives.[17] Positive correlation means that the growth of the mass of players playing each action $\dot{x}_a^p$ should have positive correlation with its current payoff $F_a^p(\mathbf{x})$, i.e. $\mathbf{F}(\mathbf{x}) \cdot \dot{\mathbf{x}} \geq 0$, and especially that $\mathbf{F}(\mathbf{x}) \cdot \dot{\mathbf{x}} = 0$ iff $\mathbf{x}$ is a Nash equilibrium. Geometrically, the transition vector $\dot{\mathbf{x}}$ from a state $\mathbf{x}$ always forms an acute angle with the payoff vector $\mathbf{F}(\mathbf{x})$ at the state.

Under the BRD, $\mathbf{F}^p(\mathbf{x}) \cdot \dot{\mathbf{x}}$ is always equal to the relative payoff of the optimal action: since $\dot{\mathbf{x}}^p = \mathbf{y}^p - \mathbf{x}^p$ with some $\mathbf{y}^p \in B^p(\mathbf{x})$, we have

$$\mathbf{F}^p(\mathbf{x}) \cdot \dot{\mathbf{x}} = \mathbf{F}^p(\mathbf{x}) \cdot (\mathbf{y}^p - \mathbf{x}^p) = F_*^p(\mathbf{x}) - \bar{F}^p(\mathbf{x}).$$

This is always non-negative, and zero iff $\mathbf{x}^p \in B^p(\mathbf{x})$. Under the tBRD, $\mathbf{F}^p(\mathbf{x}) \cdot \dot{\mathbf{x}}$ is equal to the sum of all payoff improvements gained by revising players; this suggests the tBRD should also satisfy positive correlation. We verify this in the following result.

**Theorem 2** (Positive correlation). *Given a state $\mathbf{x} \in \mathcal{X}$, let $\mathbf{z} \in V_Q(\mathbf{x})$ be a transition vector under tBRD. Then 1) for each population $p \in \mathcal{P}$,*

$$\mathbf{F}^p(\mathbf{x}) \cdot \mathbf{z}^p = \sum_{a \in \mathcal{A}^p} x_a^p Q(\breve{F}_a^p(\mathbf{x}))\breve{F}_a^p(\mathbf{x}) \geq 0.$$

*Furthermore, suppose that Assumption 3 holds. Then 2) $\mathbf{F}^p(\mathbf{x}) \cdot \mathbf{z}^p > 0$ if and only if $\mathbf{z}^p \neq \mathbf{0}$. 3) If $\mathbf{F}(\mathbf{x}) \cdot \mathbf{z} > 0$ with some $\mathbf{z} \in V_Q(\mathbf{x})$, then $\mathbf{x}$ is not a Nash equilibrium. If $\mathbf{x}$ is not a Nash equilibrium, any $\mathbf{z} \in V_Q(\mathbf{x})$ satisfies $\mathbf{F}(\mathbf{x}) \cdot \mathbf{z} > 0$.*

*Proof.* See Appendix A.2. □

## 4.2 Stability of Nash equilibrium

We now show that tBRD preserves stability of Nash equilibria in some classes of games. This result shows not only generality of stability but also clarifies the incentive structure behind these games. As tBRD is a differential inclusion and allows multiple solution paths, we need to define 'stability' in terms of convergence to the rest point on *any* solution path. For examples of classes of games mentioned here, see Chapter 3 of Sandholm (2010b).

**Definition 1** (Sandholm, 2010b: Sec. 7.A.). Consider a differential inclusion $\dot{\mathbf{x}} \in V(\mathbf{x})$ defined over $\mathcal{X}$ and a closed set $A \subset \mathcal{X}$. $A$ is **Lyapunov stable** under $V$ if for any open neighborhood $O$ of $A$ there exists a neighborhood $O'$ of $A$ such that every solution $\{\mathbf{x}^t\}$ that starts from $O'$ remains in $O$. $A$ is **attracting** if there is a neighborhood $B$ of $A$ such that every solution that starts in $B$ converges to $A$. $A$ is globally attracting if it is attracting with $B = \mathcal{X}$. $A$ is **asymptotically stable** if it is Lyapunov stable and attracting; it is globally asymptotically stable if it is Lyapunov stable and globally attracting.

The stability of Nash equilibria in a potential game is a straightforward implication of positive correlation and is known to be useful in implementing a social optimum through a dynamic version of Pigouvian taxing.[18] A population game $\mathbf{F} : \mathcal{X} \to \mathbb{R}^A$ is called a **potential game** if there is a real continuously differentiable function $f : \mathcal{X} \to \mathbb{R}$ whose gradient vector always coincides with the relative payoff vector: for all $p \in \mathcal{P}$ and $\mathbf{x} \in \mathcal{X}$

$$\frac{\partial f}{\partial x_a^p}(\mathbf{x}) = F_a^p(\mathbf{x}) - \bar{F}^p(\mathbf{x}) \text{ for all } a \in \mathcal{A}^p,$$

---

[17] See Sandholm (2010b, Sec.5.3). Among major dynamics, the logit dynamic does not satisfy Nash stationarity or positive correlation. The replicator dynamic does not satisfy Nash stationarity, as it may have a rest point that is not a Nash equilibrium.

[18] See Sandholm (2002, 2005).

$$\text{i.e. } \nabla^p f(\mathbf{x}) := \left( \frac{\partial f}{\partial x_1^p}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_{A^p}^p}(\mathbf{x}) \right)^T = \mathbf{F}^p(\mathbf{x}) - \bar{F}^p(\mathbf{x})\mathbf{1}.$$

The class of potential games includes random matching in symmetric games, binary choice games and a standard congestion games. The *potential* function $f$ works as a Lyapunov function in a wide range of evolutionary dynamics: replicator, BRD, etc. So in a potential game, the set of Nash equilibria is globally asymptotically stable under these dynamics. Now we add the tBRD to the list of stable dynamics in potential games.

**Theorem 3.** *Consider a potential game* $\mathbf{F} : \mathcal{X} \to \mathbb{R}^A$ *with a twice continuously differentiable potential function* $f$.[19] *Then, the set of Nash equilibria* $NE(\mathbf{F})$ *is globally attracting. Moreover, each local maximizer of* $f$ *is Lyapunov stable under any tBRD with Assumptions 2 and 3.*

*Proof.* From the definition of a potential function and the fact that $\mathbf{1} \cdot \dot{\mathbf{x}}^p = 0$, possitive correlation implies $\dot{f}(\mathbf{x}) = \sum_p \nabla^p f(\mathbf{x}) \cdot \dot{\mathbf{x}}^p \geq 0$, especially $\dot{f}(\mathbf{x}) = 0$ iff $\mathbf{x} \in NE(\mathbf{F})$. So $f$ is a strict Lyapunov function. Then, each local maximizer of $f$ is Lyapunov stable and the set of stationary points, i.e. $NE(\mathbf{F})$, is globally attracting (Sandholm, 2010b, Theorems 7.B.2,4). $\qquad\square$

The existence of a potential function seems to be a strong assumption on a game. Stable games are a generalization of potential games with concave potential functions. A population game $\mathbf{F}$ is a **stable game** if

$$(\mathbf{y} - \mathbf{x}) \cdot (\mathbf{F}(\mathbf{y}) - \mathbf{F}(\mathbf{x})) \leq 0 \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

The class of stable games includes two-player zero-sum games as well as games with an interior evolutionary stable state or neutrally stable state.

Hofbauer and Sandholm (2009) show that the set of Nash equilibria of a stable game is globally asymptotic stable under a broad class of evolutionary dynamics. Unlike a potential game, a stable game requires us to find a Lyapunov function by ourselves. Looking carefully at known Lyapunov functions in other dynamics, we can find that the basic idea seems to track the increase of *net* payoff improvement *of revising agents*, whose meaning depends on the revision protocol of a dynamic.[20] Seeing $Q$ as the distribution of stochastic status-quo biases added to the genuine payoff rather than exogenous tempering of revision rates, we indeed obtain a Lyapunov function $L(\mathbf{x})$ as the sum of all revising players' net payoff increases $\breve{F}_a^p(\mathbf{x}) - q$ integrated over the possible status-quo biases $q$:

$$L(\mathbf{x}) := \sum_{p \in \mathcal{P}} \sum_{a \in \mathcal{A}^p} x_a^p \int_0^{\breve{F}_a^p(\mathbf{x})} (\breve{F}_a^p(\mathbf{x}) - q) Q'(q) dq \tag{6}$$

**Theorem 4.** *Consider a stable game with Assumption 1. Then, the set of Nash equilibria* $NE(\mathbf{F})$ *is globally asymptotically stable under any tBRD with Assumptions 2 and 3.*

*Proof.* See Appendix A.3 $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

---

[19]Note that the *twice* continuous differentiability of $f$ is imposed for the continuous differentiability of $\mathbf{F} = \nabla f$, i.e. Assumption 1.

[20]For the Lyapunov functions in other dynamics, see Hofbauer and Sandholm (2009). In the BRD $\dot{\mathbf{x}} \in B(\mathbf{x}) - \mathbf{x}$, it is the difference between the optimized payoff and the current average payoff: $L(\mathbf{x}) = \sum_p F_*^p(\mathbf{x}) = \max_{\mathbf{y} \in \mathcal{X}} (\mathbf{y} - \mathbf{x})' \mathbf{F}(\mathbf{x})$. In the perturbed BRD $\dot{\mathbf{x}} \in \tilde{B}(\mathbf{x}) - \mathbf{x}$ with $\tilde{B}^p(\mathbf{x}) = \arg\max_{\mathbf{y} \in \mathcal{X}^p} \mathbf{y}' \mathbf{F}^p(\mathbf{x}) - v^p(\mathbf{y}^p)$, it is the difference between the maximized payoff and the average payoff net of the payoff perturbations: $L(\mathbf{x}) = \sum_p \left[ \max_{\mathbf{y} \in \mathcal{X}^p} (\mathbf{y} - \mathbf{x})' \mathbf{F}^p(\mathbf{x}) - (v^p(\mathbf{y}^p) - v^p(\mathbf{x}^p)) \right]$. Note that, in these dynamics, the rate at which revision opportunities are received is independent of the current payoff; thus, the average payoff of the population is equal to that of the revising agents, while they are different in tBRD because of its payoff-sensitive revision rates.

12

Furthermore, we can extend this idea of a Lyapunov function to local stability of a 'regular' ESS, by modifying it for a boundary equilibrium. A state $\mathbf{x}^* \in \mathcal{X}$ is a **regular (Taylor) evolutionary stable state** if it is a quasi-strict equilibrium[21] and it satisfies

$$(\mathbf{y} - \mathbf{x}^*) \cdot D\mathbf{F}(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}^*) < 0 \text{ whenever } (\mathbf{y} - \mathbf{x}^*) \cdot \mathbf{F}(\mathbf{x}^*) = 0 \text{ and } \mathbf{y} \in \mathcal{X} \setminus \{\mathbf{x}^*\}.$$

Let $U^p$ be the set of population $p$'s unused actions in the regular ESS $\mathbf{x}^*$. Then the latter condition can be replaced with

$$\mathbf{z} \cdot D\mathbf{F}(\mathbf{x}^*)\mathbf{z} < 0 \text{ whenever } \mathbf{z} \in T\mathcal{X} \text{ and } z_b^p = 0 \text{ for any } b \in U^p.$$

This condition means that the game $\mathbf{F}$ is a strictly stable game *locally* around the quasi-strict equilibrium $\mathbf{x}^*$ in the *reduced* state space where any action unused in $\mathbf{x}^*$ is kept unused.

For the BRD, Sandholm (2010a) constructs a Lyapunov function for a regular ESS by adding (a constant times) the total number of players who use the actions unused at the regular ESS to the Lyapunov function for a stable game. The same idea is applied to the tBRD; define a function $L^* : \mathcal{X} \to \mathbb{R}$ by

$$L^*(\mathbf{x}) = L(\mathbf{x}) + C \sum_{p \in \mathcal{P}} \sum_{b \in U^p} x_b^p,$$

where $L : \mathcal{X} \to \mathbb{R}$ is the Lyapunov function (6) for a stable game and $C \in \mathbb{R}$ is a constant. In Appendix A.4, we prove that this function $L^*$ works as a Lyapunov function for the regular ESS $\mathbf{x}^*$ when $C$ is sufficiently large positive.

**Theorem 5.** *Consider a game with Assumption 1. Suppose $\mathbf{x}^* \in \mathcal{X}$ is a regular ESS. Then, it is locally asymptotically stable under any tBRD with Assumptions 2 and 3.*

*Proof.* See Appendix A.4 □

All the stability results so far are guaranteed for all of the major evolutionary dynamics such as the best response and replicator dynamics. However, Kojima and Takahashi (2007) found a fairly general class of games where a Nash equilibrium is unique and stable in the best response and perfect foresight dynamics but the replicator and perturbed best response (e.g. logit) dynamics do not guarantee even local stability of Nash equilibrium. Single-population random matching in a symmetric two-player normal-form game is called an **anti-coordination game** if at any social state $\mathbf{x}$, the worst response is in the support of $\mathbf{x}$: namely each of the worst actions are taken by a positive mass of agents.

Kojima (2009) defines the "best response property" of a deterministic evolutionary dynamic as follows:

**Definition 2** (Kojima, 2009). An evolutionary dynamic is said to satisfy the **best response property** if any solution path $\{\mathbf{x}^t\}_{t \geq 0}$ satisfies

$$a \notin b^p(\mathbf{x}^t), x_a^{t,p} > 0 \implies \dot{x}_a^{t,p} < 0. \tag{7}$$

The BR property means that a suboptimal action never has the mass of its players increase. Perturbed BRDs such as the logit dynamic and imitative dynamics such as the replicator dynamic do not satisfy this property. It is easy to see that Assumption 3 implies that the tBRD satisfies the best response property. Kojima (2009) further verifies that this property implies global stability of the unique Nash equilibrium in an anti-coordination game. So the BR property and thus Nash stability in anti-coordination games distinguish evolutionary dynamics based on optimization like the BRD and the tBRD from other evolutionary dynamics like the replicator dynamic.

---

[21]$\mathbf{x}^*$ is a quasi-strict equilibrium, if $F_*(\mathbf{x}^*) = F_a(\mathbf{x}^*) > F_b(\mathbf{x}^*)$ for any population $p \in \mathcal{P}$, any used action $a$ and any unused action $b$, i.e. whenever $x_a^* > 0$ and $x_b^* = 0$.

# 5   Payoff monotonicity of the tBRD

Payoff sensitivity of revision rates creates a clear distinction between the tempered BRD and the standard BRD. While all suboptimal actions decrease their mass of players at the same rate in the BRD, a better suboptimal action decays slower than worse ones in the tBRD. Furthermore, while an optimal action may see the mass of its players decrease just as fast as suboptimal actions do under the BRD when there are multiple best responses, Assumption 3 prevents its mass of players from decreasing under the tBRD. In sum, the growth rate of each action's players in the tBRD respects the payoff ordering and thus the tBRD is a payoff monotone selection:

**Definition 3.** An evolutionary dynamic is said to satisfy **payoff monotonicity** and is called a payoff monotone selection if any interior solution path $\{\mathbf{x}^t\}_{t \geq 0} \in \mathring{\mathcal{X}} := \mathcal{X} \cap (0,1)^A$ satisfies both of the following two conditions for almost all time $t \in \mathbb{R}_+ := [0, \infty)$, any $p \in \mathcal{P}$ and any $a, b \in \mathcal{A}^p$:[22]

$$F_a^p(\mathbf{x}^t) > F_b^p(\mathbf{x}^t) \;\Rightarrow\; \frac{\dot{x}_a^{t,p}}{x_a^{t,p}} > \frac{\dot{x}_b^{t,p}}{x_b^{t,p}}, \tag{8}$$

$$a \in b^p(\mathbf{x}^t) \;\Rightarrow\; \dot{x}_a^{t,p} \geq 0. \tag{9}$$

Furthermore, we call it a **regular** payoff monotone selection if

$$\liminf_{t \to \infty} \{ F_a^p(\mathbf{x}^t) - F_b^p(\mathbf{x}^t) \} > 0 \;\Rightarrow\; \liminf_{t \to \infty} \left( \frac{\dot{x}_a^{p,t}}{x_a^{p,t}} - \frac{\dot{x}_b^{p,t}}{x_b^{p,t}} \right) > 0.$$

The replicator dynamic, perturbed best response dynamics such as the logit dynamic, and the tBRD are regular payoff monotone selections, while the standard BRD and its recent variants such as the refined BRD and the sampling BRD are not.[23] The tBRD resembles the replicator dynamic in some games because of payoff monotonicity, while it behaves more like the BRD than the replicator dynamic in others, e.g. anticoordination games, because of the BR property.[24] In the following, we see cases where the payoff monotonicity of the tBRD eliminates pathological outcomes from the standard BRD.

## 5.1   Elimination of pathological indeterminacy and cycles

In the example below, the BRD allows multiple solution paths from the interior Nash equilibrium, including cycles and escaping paths to other Nash equilibria (Hofbauer, 1995, Example 3.2). As such multiplicity takes place in a positive-measure set of initial states, this multiplicity cannot be negligible.

The payoff monotonicity of tBRD determines a unique Carathèodory solution path from all points; in particular, it selects only the stationary path from the interior equilibrium, and retains local stability of the interior Nash equilibrium like replicator dynamic.

*Example* 3 (Zeeman's game). Consider a single-population random matching game with $\mathcal{A} = \{1, 2, 3\}$ and the payoff function $\mathbf{F}$ given by

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} 0 & 6 & -4 \\ -3 & 0 & 5 \\ -1 & 3 & 0 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 6x_2 - 4x_3 \\ -3x_1 + 5x_3 \\ -x_1 + 3x_2 \end{pmatrix}. \tag{10}$$

---

[22]In the preceding literature (e.g. Weibull (1995, Definition 4.2.) and Hofbauer and Sigmund (1998, p.88)), payoff monotonicity requires "two-sided" monotonicity: $F_a^p(\mathbf{x}^t) > F_b^p(\mathbf{x}^t) \Leftrightarrow \dot{x}_a^{t,p}/x_a^{t,p} > \dot{x}_b^{t,p}/x_b^{t,p}$, which implies both of (8) and (9). The tBRD does not satisfy the two-sided monotonicity, because multiple optimal actions can have the masses of their players grow at different rates. See Zusai (2012b).

[23]See Balkenborg, Hofbauer, and Kuzmics (2012) for the refined BRD and Oyama, Sandholm, and Tercieux (2010) for the sampling BRD.

[24]Golman and Page (2010) present several games with multiple equilibria where the best response and replicator dynamics may yield significantly different basin of attraction sizes across equilibria of the same game. The basin of attraction under the tBRD is the same as that under the BRD in one of their examples (the Haruvy-Stahl game) and is similar to that under the replicator dynamic in another example (the game used to prove their Theorem 2).
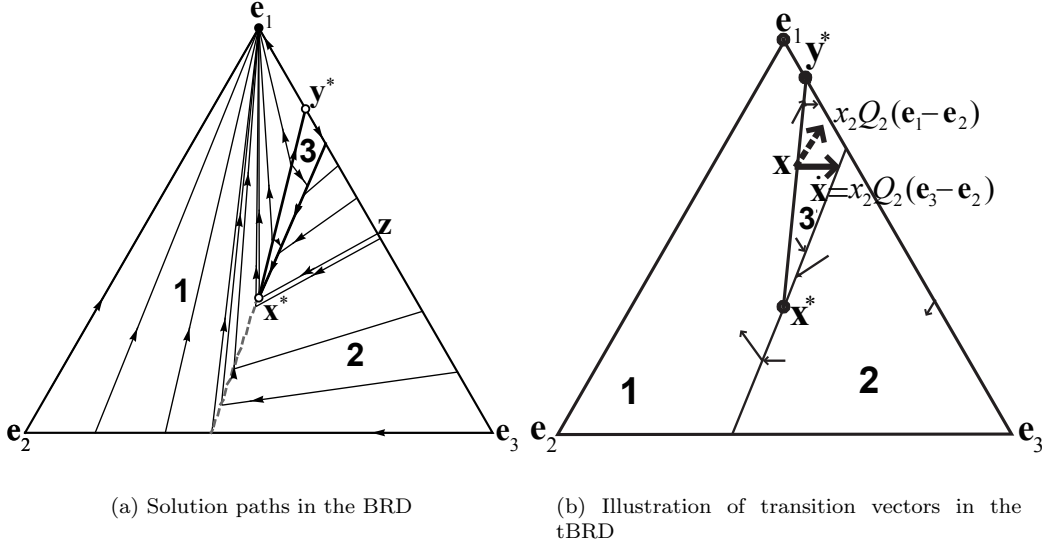
(a) Solution paths in the BRD

(b) Illustration of transition vectors in the tBRD

*Figure 3*: Zeeman's game (Example 3). In (b), $\mathbf{x}$ is near above the line $\mathbf{x}^*\mathbf{y}^*$. The number in bold shows the optimal action in the region, and the solid lines indicate boundaries of best response regions.

This game has three Nash equilibria: $\mathbf{e}_1 = (1, 0, 0), \mathbf{x}^* = (1/3, 1/3, 1/3), \mathbf{y}^* = (4/5, 0, 1/5)$.[25]

Under the BRD, there are multiple solution paths starting from each point in the triangle $\mathbf{x}^*\mathbf{y}^*\mathbf{z}$ (Hofbauer, 1995). First, from $\mathbf{x}^*$, the state can stay at $\mathbf{x}^*$ or move towards $\mathbf{e}_1$ or towards $\mathbf{y}^*$. Second, at any point on the boundary of the BR regions of actions 2 and 3, $b^{-1}(2) \cap b^{-1}(3)$, the state can move towards $\mathbf{e}_1$ or $\mathbf{y}^*$, or even towards $\mathbf{e}_3$.[26] If it moves towards $\mathbf{e}_3$, the path hits the boundary $b^{-1}(2) \cap b^{-1}(3)$ and then slides on this line towards $\mathbf{x}^*$. After reaching $\mathbf{x}^*$ in finite time, the multiplicity of transition vectors remains. So, from any point in $b^{-1}(3)$, the BRD allows continuously many trajectories. In particular, we cannot tell whether or not the interior Nash equilibrium can be 'sustained' in the BRD.

Look carefully at the transition from the state on the boundary $b^{-1}(1) \cap b^{-1}(3)$. When it goes to $\mathbf{e}_1$ or $\mathbf{y}^*$ in the tBRD, action 1 increases its players and remains optimal. The suboptimal action 2, which favors action 1 more than action 3, decreases its players; action 1 could not remain optimal without significant decrease of action 3. But action 3 players do not have an incentive to abandon this action when the social state is on the boundary, so the BRD paths going to $\mathbf{e}_1$ or $\mathbf{y}^*$ do not have sensible economic interpretations.

The tBRD eliminates all of the sources of multiple solution paths. First, the strong Nash stationarity (part 3 of Theorem 1) eliminates any escaping paths starting from the interior Nash equilibrium $\mathbf{x}^*$. Furthermore, payoff monotonicity forces the social state on the BR boundary $b^{-1}(1) \cap b^{-1}(3)$ to enter $b^{-1}(3)$. Under the tBRD, only action 2 players revise their actions on this boundary, so action 1 cannot maintain its optimality and action 3 becomes the unique best response. We can geometrically determine the transition vectors near each boundary of $b^{-1}(\cdot)$ from such an argument; see Fig. 3 (b).

Fig. 4 illustrates that, starting from any point sufficiently close to the interior Nash equilibrium $\mathbf{x}^*$, the social state converges to $\mathbf{x}^*$ under the tBRD: $\mathbf{x}^*$ is an *attracting* rest point. There is no deviation from the interior Nash equilibrium $\mathbf{x}^*$ to other Nash equilibria $\mathbf{y}^*$ and $\mathbf{e}_1$ under the tBRD. In particular, any cycle starting from $\mathbf{x}^*$ or from $b^{-1}(3)$ is impossible. The social state may, however, escape from a neighborhood of $\mathbf{x}^*$ for a while,

---

[25]Originally this example is presented in Zeeman (1980) to show that an interior Nash equilibrium that is not an evolutionary stable state (ESS) can be asymptotically stable under the replicator dynamic. See also Sandholm (2010b, Example 5.1.7).

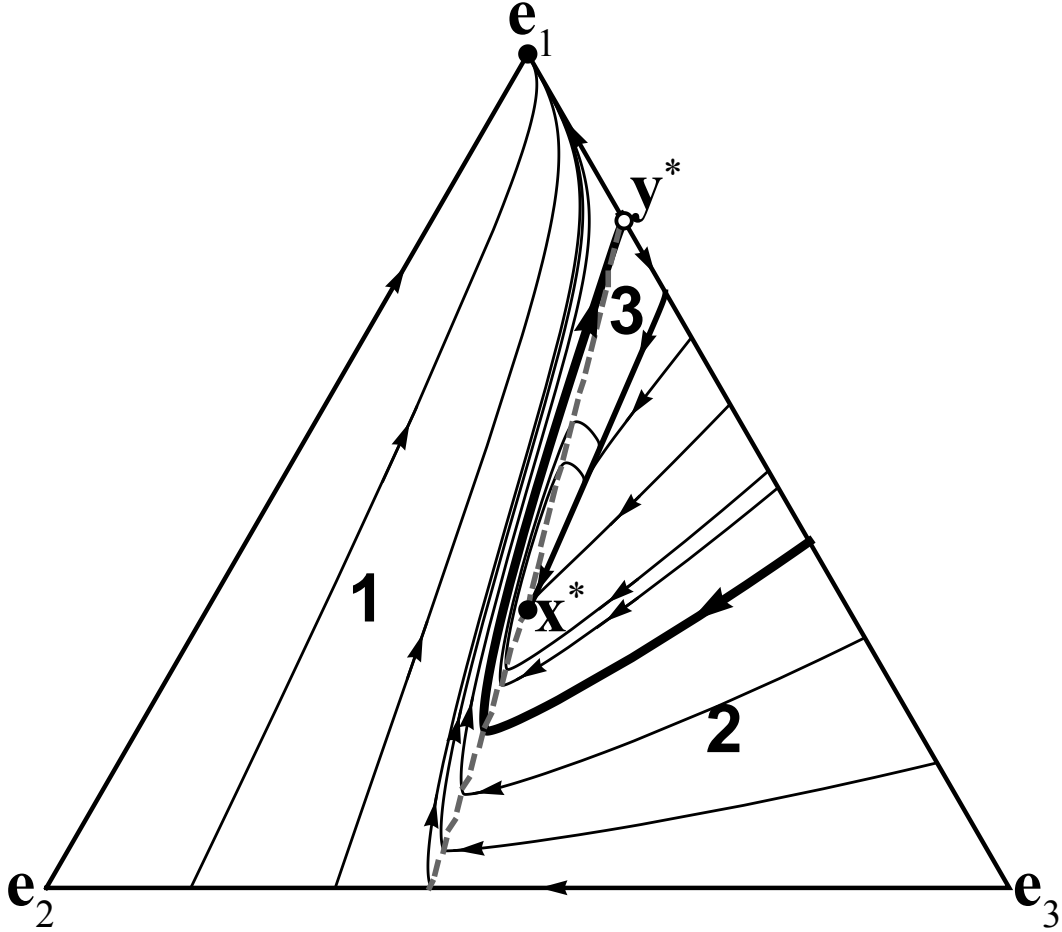[26]Note that the transition vectors to $\mathbf{e}_1$ and $\mathbf{e}_3$ form an obtuse angle.

*Figure 4*: Solution paths under the tBRD in Zeeman's game (Example 3), drawn with Dynamo. The boldest line ending at $\mathbf{y}^*$ separates the basins of attraction to the Nash equilibria $\mathbf{e}_1$ and $\mathbf{x}^*$.

as seen in Fig. 4; so $\mathbf{x}^*$ is not Lyapunov stable under the tBRD, unlike under the replicator dynamic, where $\mathbf{x}^*$ is asymptotically stable, namely Lyapunov stable as well as attracting.
▌

## 5.2 Interior convergence and equilibrium refinement

It is a natural idea to see an interior convergence path under an evolutionary dynamic as a sequence of perturbed states to define a refined equilibrium. Hofbauer (1995) proves that a Nash equilibrium is trembling-hand perfect in a normal form game if and only if there is an interior convergence path to it under the BRD in a random matching of the game. Zusai (2012b) investigates a situation where such a connection holds between regular payoff monotone selections and proper equilibrium.[27] Since the tBRD is a regular payoff monotone selection, we can apply this result to the tBRD.

---

[27] A social state $\mathbf{x} \in \mathcal{X}$ is an $\varepsilon$-proper equilibrium with $\varepsilon > 0$, if $\mathbf{x}$ lies in the interior of $\mathcal{X}$ and it satisfies $x_a^p < \varepsilon x_b^p$ whenever $F_a^p(\mathbf{x}) < F_b^p(\mathbf{x})$ for all $p \in \mathcal{P}, a, b \in \mathcal{A}^p$. A social state $\mathbf{x}^* \in \mathcal{X}$ is a proper equilibrium if there are sequences $\{\mathbf{x}^n\} \subset \mathcal{X}$ and $\{\varepsilon^n\} \subset (0, \infty)$ such that each $\mathbf{x}^n$ is an $\varepsilon^n$-proper equilibrium and the sequence $\{(\mathbf{x}^n, \varepsilon^n)\}$ satisfies $\mathbf{x}^n \to \mathbf{x}^*$ and $\varepsilon^n \to 0$ as $n \to \infty$.

In a simple two-stage chain-store game, an interior path converges to a Nash equilibrium with a weakly dominated strategy both under the replicator dynamic and under the tBRD, while only to a strict equilibrium under the BRD. (Zusai, 2012b; Cressman, 2003, p.291.) So this connection cannot be generalized.

**Theorem 6.** *Consider a tBRD $\dot{\mathbf{x}} \in V_Q(\mathbf{x})$ in a random matching of a normal form game $G$. Assume Assumptions 1–3. Suppose that there exists an interior path $\{\mathbf{x}^t\}_{t \in \mathbb{R}_+} \subset \mathring{\mathcal{X}}$ converging to a state $\mathbf{x}^\infty \in \mathcal{X}$ under the tBRD.*

*Furthermore, suppose that (i) $sgn(F_a^p(\mathbf{x}^t) - F_b^p(\mathbf{x}^t)) = sgn(F_a^p(\mathbf{x}^0) - F_b^p(\mathbf{x}^0))$ for all $t \in \mathbb{R}_+$ and $a, b \in \mathcal{A}^p, p \in \mathcal{P}$, (ii-1) $x_a^{\infty,p} = 0$ if $F_a^p(\mathbf{x}^t) < F_*^p(\mathbf{x}^t)$ at any $t \in \mathbb{R}_+$ and $a \in \mathcal{A}^p$, and (ii-2) $F_a^p(\mathbf{x}^\infty) = F_b^p(\mathbf{x}^\infty)$ implies $F_a^p(\mathbf{x}^t) = F_b^p(\mathbf{x}^t)$ for all $t \in \mathbb{R}_+$ unless either $a$ or $b$ is optimal on the path. Then $\mathbf{x}^\infty$ is a proper equilibrium in $G$. If $G$ is the normal form of an extensive-form game, $\mathbf{x}^\infty$ is a sequential equilibrium and thus a subgame-perfect equilibrium.*

*Proof.* Zusai (2012b) verifies this theorem for regular payoff monotone selections. We apply it to the tBRD, as it is a regular payoff monotone selection.  □

Proper equilibrium is stronger than trembling-hand perfect equilibrium and implies sub-game perfect equilibrium in any corresponding extensive form, while trembling-hand perfect equilibrium does not. In the following example, the tBRD has a stronger power to select equilibrium than the BRD, because only a proper equilibrium is chosen as a stable state in the tBRD. Note that equilibrium refinements such as proper equilibrium require agents' beliefs to be so sophisticated as to take others' rationality into account.[28] The result here suggests that myopic interaction between boundedly rational agents eventually reaches the same outcome in the long run.

*Example* 4. Consider a two-player sequential-move game in Fig. 5 (van Damme, 1991, Fig. 6.5.1). In this example, both $(A, r)$ and $(L, l)$ are trembling-hand perfect in the normal form and sequential equilibria in the extensive form. $(L, l)$ is the only proper equilibrium and $(A, r)$ is not. Actually, $(A, r)$ is not a plausible outcome, because to choose $r$, player 2 should believe that player 1 plays $R$ with higher probability than $L$. But $R$ is always worse than $L$ for player 1 and thus this belief seems inconsistent with 1's rationality. We wonder if such an implausible outcome would be sustained in the long run.
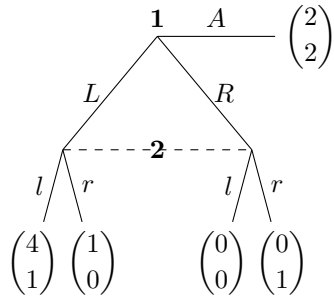
Let us consider two-population random matching in the normal form of this game. Under the BRD, since $(A, r)$ is perfect, there is an interior convergence path to it. Actually, if the social state lies in the interior of $b^{-1}(A, r) = \{\mathbf{x} \in \mathcal{X} | x_L^1 \leq x_R^1, x_l^2 \leq 1/3\}$, the unique best response is $(A, r)$ and the social state linearly converges to $(\mathbf{e}_A^1, \mathbf{e}_r^2)$. But this path does not respect payoff monotonicity: $R$ is worse than $L$ but both decrease their players at the same rates on this path.

Under the tBRD, payoff monotonicity eliminates such paths. Consider an interior solution path starting from $b^{-1}(A, r)$ under the tBRD. Theorem 6 guarantees that the state does not reach the non-proper equilibrium $(\mathbf{e}_A^1, \mathbf{e}_r^2)$ without reversing either inequality $x_L^1 \leq x_R^1$ or $x_l^2 \leq 1/3$. As $x_l^2$ decreases and the latter inequality keeps holding as long as the former inequality is satisfied, the former should be reversed and the state should escape from $b^{-1}(A, r)$ to $b^{-1}(A, l)$. (See period 19 in Figure 5 (d,f).) After this, $l$ remains optimal because payoff monotonicity keeps the inequality $x_L^1 \geq x_R^1$. $x_l^2$ eventually exceeds $1/3$ and the social state goes into $b^{-1}(L, l)$. (See period 48.) Then, $L$ becomes optimal for population 1, as well as $l$ for population 2. So the social state converges to $(\mathbf{e}_L^1, \mathbf{e}_l^2)$, i.e. the proper equilibrium $(L, l)$. This convergence result is readily generalized to any regular monotone selection; see Zusai (2012b). ▮

# 6  Continuous payoff-sensitivity of the tBRD

One of the disadvantages of the standard BRD is discontinuity of the transition vector around equilibrium, which disables linear approximation and local stability analysis by the Hessian matrix. The continuous payoff sensitivity of the tBRD reduces the prevalence of this kind of discontinuity, though the transition vector remains discontinuous on the boundary of best response regions except at equilibrium. Here we argue that this allows linear approximation of the tBRD.

---

[28]For epistemological foundation of properness, see Blume, Brandenburger, and Dekel (1991).

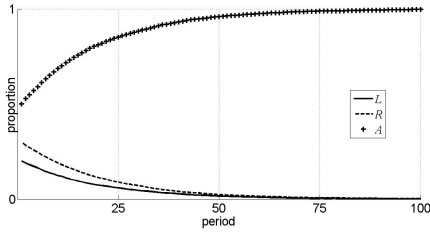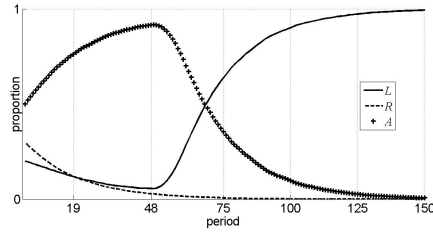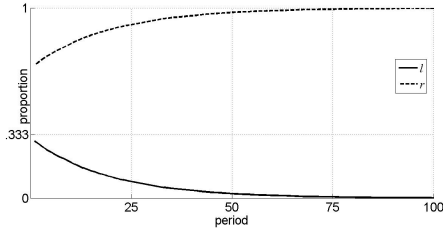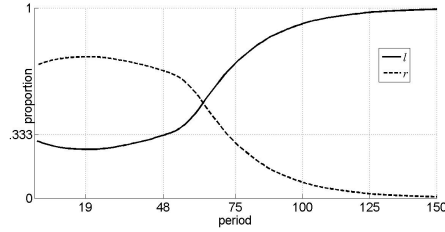|   | 2 | $l$ | $r$ |   |
|---|---|-----|-----|---|
| 1 |   |   |   |   |
| $L$ |   | 1 ⟍ 4 | 0 ⟍ 1 | $x_l^2 > 1/3$ |
| $R$ |   | 0 ⟍ 0 | 1 ⟍ 0 | always worst |
| $A$ |   | 2 ⟍ 2 | 2 ⟍ 2 | $x_l^2 < 1/3$ |
|   |   | $x_L^1 > x_R^1$ | $x_L^1 < x_R^1$ |   |

(a) Extensive form

(b) Normal form

(c) $\mathbf{x}^1$ in BRD

(d) $\mathbf{x}^1$ in tBRD

(e) $\mathbf{x}^2$ in BRD

(f) $\mathbf{x}^2$ in tBRD

*Figure 5*: Example 4. The inequality out of each column/row of the table is the condition for the action in this column/row to be the best response. The width of an arrow shows the length of the transition vector, while the length of an arrow is normalized. The graphs (c)–(f) are obtained from discrete-time finite-population simulations.

Thanks to continuity (Assumption 2) of the tempering function $Q$, the speed of transition $\|\dot{\mathbf{x}}\|$ changes continuously under the tBRD, while it does not udner the BRD, when the state leaves an equilibrium. When the state escapes from a mixed equilibrium $\mathbf{x}^*$ toward one of the pure best responses, say action 1, the speed discontinuously grows from zero to $\|\mathbf{e}_1 - \mathbf{x}^*\| \gg 0$ under the BRD; on the other hand, it gradually increases from zero under the tBRD. In particular, such a sudden change creates a difference between the BRD and the tBRD with regard to strong Nash stationarity (part 3 of Theorem 1).

Because of its continuity and strong stationarity, the tBRD can be seen as a system of piecewise differential equations (piecewise DEs), also known as a switched system in engineering. In a system of piecewise DEs, the state space is partitioned into several almost disjoint regions $\mathcal{X}_i$ where a single differential equation $\dot{\mathbf{x}} = V_i(\mathbf{x})$ with a continuous function $V_i : \mathcal{X}_i \to T\mathcal{X}$ governs the motion of the state. On a (measure-zero) intersection of such regions, any convex combination of transition vectors that are feasible in the neighbor

18

regions are allowed.[29] A rest point can be on the intersection, but it must be stationary in *each* differential equation. To investigate dynamic properties in a system of piecewise DEs, we decompose the entire system to those of each differential equation and then impose some additional conditions to mitigate discontinuity caused by switches between differential equations.

Both under the BRD and under the tBRD, once we focus on (the interior of) a best response region of a single action profile $\mathbf{b} \in \mathcal{A}$, we obtain a differential equation such as $\dot{\mathbf{x}}^p = \mathbf{e}_{b^p}^p - \mathbf{x}^p$ for the BRD and $\dot{\mathbf{x}}^p = \sum_a x_a^p Q(\check{F}_a^p(\mathbf{x}^p))(\mathbf{e}_{b^p}^p - \mathbf{e}_a^p)$ for the tBRD. In the former, a mixed Nash equilibrium cannot be a rest point under the decomposed dynamic defined on the single BR region. On the other hand, it is still a rest point under the decomposed dynamic of the tBRD because $x_a^p Q(\check{F}_a^p(\mathbf{x})) = 0$ for all $a, p$ iff $\mathbf{x}$ is a Nash equilibrium. So we can redefine the tBRD as a system of piecewise DEs such as

$$\dot{\mathbf{x}}^p = \sum_{a^p \in \mathcal{A}^p} x_{a^p}^p Q(F_{b^p}^p(\mathbf{x}) - F_{a^p}^p(\mathbf{x}))(\mathbf{e}_{b^p}^p - \mathbf{e}_{a^p}^p) =: V_Q^p(\mathbf{x}; \mathbf{b}) \quad \text{for all } \mathbf{x} \in b^{-1}(\mathbf{b}) \qquad (11)$$

for each $\mathbf{b} := (b^1, \cdots, b^P) \in \mathcal{A}$. Note that the tBRD allows the group of revising agents to take any convex combination of pure best responses in a neighborhood of the current state. To have the BR regions as almost disjoint regions, we assume the following:

**Assumption 4.** The best response correspondence $b : \mathcal{X} \rightrightarrows \mathcal{A}$ is single-valued almost everywhere in $\mathcal{X}$ and every Nash equilibrium is isolated.

We can apply findings in the literature on piecewise DEs to the tBRD. Interestingly, analysis of piecewise DEs was applied to economics in early 1980's to study non-Walrasian disequilibrium macroeconomic dynamics. Here we import some results from Honkapohja and Ito (1983) to the tBRD.

**Theorem 7.** *Consider a tBRD as a system of piecewise DEs* (11). *Assume Assumptions 1–4. Consider any* $\mathbf{b}, \mathbf{c} \in \mathcal{A}$ *whose BR regions have non-empty intersections, i.e.* $b^{-1}(\mathbf{b}) \cap b^{-1}(\mathbf{c}) \neq \emptyset$. *Then, suppose that the following two conditions hold in* $b^{-1}(\mathbf{b}) \cap b^{-1}(\mathbf{c})$ *for some population p with* $b^p \neq c^p$.

*i) (a) The equation* $F_{b^p}^p(\mathbf{x}) - F_{c^p}^p(\mathbf{x}) = 0$ *is solvable for either one coordinate of* $\mathbf{x}$, *and (b) the solution is* $C^2$-*class;*

*ii) (a)* $(DF_{b^p}^p(\mathbf{x}) - DF_{c^p}^p(\mathbf{x}))V_Q(\mathbf{x}; \mathbf{b}) < 0$ *or (b)* $(DF_{b^p}^p(\mathbf{x}) - DF_{c^p}^p(\mathbf{x}))V_Q(\mathbf{x}; \mathbf{c}) > 0$.

*Then, there exists a unique solution path from each point in* $\mathcal{X}$.

*Proof.* Note that the uniqueness of a solution path is guaranteed in the interior of each BR region, where only one single differential equation governs the transition.

By Assumption 1 and 2, $V_Q(\cdot; \mathbf{b}')$ is $C^1$ with any $\mathbf{b}' \in \mathcal{A}$. (i) suggests that the function $F_{b^p}^p - F_{c^p}^p$ separates $b^{-1}(\mathbf{b})$ and $b^{-1}(\mathbf{c})$: $\mathbf{x}$ belongs to the interior of $b^{-1}(\mathbf{b})$ if it is positive, $b^{-1}(\mathbf{b}) \cap b^{-1}(\mathbf{c})$ if zero, and the interior of $b^{-1}(\mathbf{c})$ if negative. Then, we can apply Honkapohja and Ito (1983, Thm.2.1) to the uniqueness of a solution path under the tBRD around the boundaries of the BR regions. □

The above conditions are imposed for the uniqueness of the transition vector when the state crosses the border of two regions. Condition ii) means that the state is pushed out from one region to the other region; if both (ii-a) and (ii-b) are satisfied, the state is forced to slide on the boundary. From Fig. 3 (b), we easily find Zeeman's game (Example 3) satisfying condition ii), as well as i) by the linearity of the payoff function and the non-singularity of its payoff matrix. So we assure the uniqueness of the solution in this game.

---

[29]That is, Filippov solutions are usually adopted as the solution concept for a system of piecewise DEs. Under the tBRD they coincide with Carathéodory solutions because of the upper semicontinuity and the convexity of $V_Q$, though this is not true in general.

Furthermore, we can check local stability of an isolated equilibrium from the Hessian matrices of neighbor DEs. Contrary to what might be expected, stability in each differential equation does not guarantee local stability in the whole system of piecewise DEs. We need somewhat a stronger condition so that the dynamics in separate regions are well coordinated.[30]

**Theorem 8.** *Consider a tBRD as a system of piecewise DEs* (11). *Assume Assumptions 1–4. A Nash equilibrium* $\mathbf{x}^*$ *is locally asymptotically stable if*

$$\mathbf{z} \cdot \{DV_Q(\mathbf{x}^*; \mathbf{b}) + (DV_Q(\mathbf{x}^*; \mathbf{b}))^T\}\mathbf{z} < 0 \text{ for any } \mathbf{z} \in T\mathcal{X} \setminus \{\mathbf{0}\}$$

*for each* $\mathbf{b} \in b(\mathbf{x}^*)$.

*Proof.* This is a straightforward application of Honkapohja and Ito (1983, Cor.2). □

# 7  Concluding remarks

We defined the tempered best response dynamic by combining payoff-dependent revision rates with inertia and myopic optimization. Payoff-dependent revision rates mitigate the discontinuity of the standard best response protocol. Stability of Nash equilibrium in various classes of games is robust to this modification. Similar to the replicator dynamic, the tBRD is a payoff monotone selection. This property eliminates pathological multiplicity of solutions and implausible cycles in Zeeman's game (Example 3) and makes the long-run outcome more consistent with the prediction from equilibrium refinement than under the BRD in a simple sequential-move game (Example 4).

The speed of transition grows only continuously when the social state departs from an equilibrium. A mixed Nash equilibrium is a rest point even in each best response region, unlike under the standard BRD. This enables us to apply results on systems of piecewise continuous differential equations to the tBRD. In this paper, we give sufficient conditions for uniqueness of a solution path and for local stability of equilibrium. The latter allows us to investigate stability by linearization. Among optimization-based dynamics, the logit dynamic is conventionally used to analyze stability analytically, but its rest points are perturbed from Nash equilibria, depending on the noise level. Under the tBRD, rest points coincide with Nash equilibria. The tBRD may be useful for applied theorists who want to confirm dynamic stability of Nash equilibria without perturbation of rest points.

In a game with heterogeneous payoff types, the payoff dependency of revision rates in the tBRD endogenously produces heterogeneity in revision rates among different payoff types. Zusai (2012a) applies the tBRD to a binary Bayesian game with heterogeneous payoff types and considers the implications for dynamic implementation of social optima.

# A  The proofs

## A.1  Theorem 1 (Nash stationarity)

*Proof.* 1) Suppose $\mathbf{x}$ is a Nash equilibrium. Then for each $p \in \mathcal{P}$, every action $a \in \mathcal{A}^p$ satisfies $x_a^p(\mathbf{x}) = 0$ or $F_a^p(\mathbf{x}) = F_*^p(\mathbf{x})$. By Assumption 3, the latter implies $Q(\check{F}_a^p(\mathbf{x})) = 0$. In both cases, (almost) every action-$a$ player does not revise his action: $x_a^p(\mathbf{x})Q(\check{F}_a^p(\mathbf{x})) = 0$. So we obtain $V(\mathbf{x}) = \{\mathbf{0}\}$ by (2).

2) Suppose $\mathbf{x}$ is not a Nash equilibrium. Then we can find at least one population $p \in \mathcal{P}$ with a suboptimal action $a \in \mathcal{A}^p$ being played by a positive mass of its players $x_a^p > 0$. Like under the BRD, under the tBRD this mass decreases at the speed $x_a^p Q(\check{F}_a^p(\mathbf{x})) > 0$. So any transition vector $\dot{\mathbf{x}} \in V(\mathbf{x})$ has a negative entry $\dot{x}_a^p(\mathbf{x}) = -x_a^p Q(\check{F}_a^p(\mathbf{x})) < 0$ for this action. Hence $\dot{\mathbf{x}}$ cannot be a zero vector, i.e. $\mathbf{0} \notin V(\mathbf{x})$.

---

[30]See Honkapohja and Ito (1983) for a weaker sufficient condition for stability.

3) First of all, Lipschitz continuity of $\mathbf{F}$ implies Lipschitz continuity of the revision incentive $\breve{F}_a^p := F_*^p - F_a^p$ for any population $p \in \mathcal{P}$ and action $a \in \mathcal{A}^p$. Let $\bar{K}^p$ be the largest Lipschitz constant of $\breve{F}_a^p$ among all $a \in \mathcal{A}^p$:

$$|\breve{F}_a^p(\mathbf{x}_1) - \breve{F}_a^p(\mathbf{x}_2)| \leq \bar{K}^p |\mathbf{x}_1 - \mathbf{x}_2| \text{ for all } \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}.$$

Suppose there is a Carathéodory solution path staying at a Nash equilibrium $\mathbf{x}^*$ until time $T \geq 0$ and leaving it at time $T$. Then by (2) and the triangle inequality we have

$$|\dot{\mathbf{x}}^t| \leq \sum_{a,p} x_a^{p,t} |Q(\breve{F}_a^p(\mathbf{x}^t))| \cdot |\mathbf{y}_a^{p,t} - \mathbf{e}_a^{p,t}| \quad \text{with some } \mathbf{y}_a^{p,t} \in B^p(\mathbf{x}^t) \subset \Delta\mathcal{A}^p$$

for almost all time $t$. First, consider any action $a \notin b^p(\mathbf{x}^*)$. Nash equilibrium requires $x_a^{p,*} = 0$ and continuity of $\mathbf{F}^p$ and the path $\mathbf{x}^t$ implies $\breve{F}_a^p(\mathbf{x}^{T+\tau}) > 0$ for sufficiently small $\tau$. So in the time range $[T, T+\tau]$, such actions are not optimal and thus each keeps $x_a^{p,t} = 0$. Second, consider an action $a \in b^p(\mathbf{x}^*)$. Then $F_*^p(\mathbf{x}^*) = F_a^p(\mathbf{x}^*)$ and thus $Q(\breve{F}_a^p(\mathbf{x}^*)) = Q(0) = 0$ by Assumption 3. With this fact, Lipschitz continuity of $Q$ and $\breve{F}_a^p$ yields

$$\begin{aligned}
|Q(\breve{F}_a^p(\mathbf{x}^t))| &= |Q(\breve{F}_a^p(\mathbf{x}^t)) - Q(\breve{F}_a^p(\mathbf{x}^*))| \\
&\leq K_Q |\breve{F}_a^p(\mathbf{x}^t) - \breve{F}_a^p(\mathbf{x}^*)| \leq K_Q \cdot 2\bar{K}^p |\mathbf{x}^t - \mathbf{x}^*| \qquad \text{for any } a \in b^p(\mathbf{x}^*).
\end{aligned}$$

Hence in either case, we have

$$x_a^{p,t} |Q(\breve{F}_a^p(\mathbf{x}^t))| \leq 2\bar{K}^p K_Q x_a^{p,t} |\mathbf{x}^t - \mathbf{x}^*| \quad \text{for any } p \in \mathcal{P}, a \in \mathcal{A}^p \text{ and all } t \in (T, T+\tau].$$

With $|\mathbf{y}^{\cdot}|, |\mathbf{e}^{\cdot}| \leq 1$ and $\sum_a x_a^{p,t} = 1$, this implies

$$|\dot{\mathbf{x}}^t| \leq \sum_{a,p} 2\bar{K}^p K_Q x_a^{p,t} |\mathbf{x}^t - \mathbf{x}^*| \cdot 2 \leq K_V |\mathbf{x}^t - \mathbf{x}^*| \quad \text{for almost all } t \in [T, T+\tau],$$

where $K_V = 4K_Q \sum_{a,p} \bar{K}^p < \infty$. Since $\mathbf{x}^t$ is a Carathéodory solution and thus absolutely continuous, we have

$$|\mathbf{x}^s - \mathbf{x}^T| = \int_T^s |\dot{\mathbf{x}}^t| dt \leq K_V \int_T^s |\mathbf{x}^t - \mathbf{x}^*| dt \quad \text{for almost all } s \in [T, T+\tau].$$

Then Gronwall's inequality implies

$$|\mathbf{x}^s - \mathbf{x}^*| \leq 0 \cdot \exp(K_V s) = 0 \text{ for all } s \in [T, T+\tau].$$

So we have $\mathbf{x}^s = \mathbf{x}^*$ during $s \in [T, T+\tau]$. This contradicts the hypothesis that $\mathbf{x}^t$ departs from $\mathbf{x}^*$ at time $T$. We therefore conclude that $\mathbf{x}^t \equiv \mathbf{x}^*$ is the only Carathéodory solution starting from $\mathbf{x}^0 = \mathbf{x}^*$. $\qquad\qquad\square$

## A.2 Theorem 2 (positive correlation)

*Proof.* 1) We begin the proof from the equality in 1). The vector $\mathbf{z}^p \in V^p(\mathbf{x})$ should be represented as

$$\mathbf{z}^p = \sum_{a \in \mathcal{A}^p} x_a^p Q(\breve{F}_a^p(\mathbf{x}))(\mathbf{y}_a^p - \mathbf{e}_a^p)$$

with a best response $\mathbf{y}_a^p \in B^p(\mathbf{x})$. So we have

$$\begin{aligned}
\mathbf{F}^p(\mathbf{x}) \cdot \mathbf{z}^p &= \mathbf{F}^p(\mathbf{x}) \cdot \left[ \sum_{a \in \mathcal{A}^p} x_a^p Q(\breve{F}_a^p(\mathbf{x}))(\mathbf{y}_a^p - \mathbf{e}_a^p) \right] = \sum_{a \in \mathcal{A}^p} x_a^p Q(\breve{F}_a^p(\mathbf{x})) \{\mathbf{F}^p(\mathbf{x}) \cdot \mathbf{y}_a^p - \mathbf{F}^p(\mathbf{x}) \cdot \mathbf{e}_a^p\} \\
&= \sum_{a \in \mathcal{A}^p} x_a^p Q(\breve{F}_a^p(\mathbf{x}))(F_*^p(\mathbf{x}) - F_a^p(\mathbf{x})) = \sum_{a \in \mathcal{A}^p} x_a^p Q(\breve{F}_a^p(\mathbf{x})) \breve{F}_a^p(\mathbf{x})
\end{aligned}$$

Since all terms in the last summation are non-negative, we have $\mathbf{F}^p(\mathbf{x}) \cdot \mathbf{z}^p \geq 0$.

2) It is clear that $\mathbf{z}^p = \mathbf{0}$ implies $\mathbf{F}^p(\mathbf{x}) \cdot \mathbf{z}^p = 0$, which is equivalent to the 'only-if' clause. Next we prove the 'if' clause: suppose $\mathbf{z}^p \neq \mathbf{0}$. Then the term $x_a^p Q(\check{F}_a^p(\mathbf{x}))$ is positive for some $a \in \mathcal{A}^p$ by (2). With Assumption 3, this implies $\check{F}_a^p(\mathbf{x}) > 0$. Hence $x_a^p Q(\check{F}_a^p(\mathbf{x}))\check{F}_a^p(\mathbf{x})$ is positive for this $a$, and thus $\mathbf{F}^p(\mathbf{x}) \cdot \mathbf{z}^p > 0$.

3) Notice that $\mathbf{F}(\mathbf{x}) \cdot \mathbf{z} = \sum \mathbf{F}^p(\mathbf{x}) \cdot \mathbf{z}^p$; so $\mathbf{F}(\mathbf{x}) \cdot \mathbf{z} > 0$ iff $\mathbf{F}^p(\mathbf{x}) \cdot \mathbf{z}^p > 0$ in some population $p$, and $\mathbf{F}(\mathbf{x}) \cdot \mathbf{z} = 0$ otherwise. We obtain 3) by combining 2) and weak Nash stationarity (Theorem 1.1-2). $\qquad\square$

## A.3    Theorem 4 (stability in stable games)

Here we prove the global asymptotic stability of Nash equilibria in stable games (Theorem 4) in three steps: First, we verify that the function $L$ in (6) is Lipschitz continuous both in state $\mathbf{x}$ and in time $t$. Second, we prove an auxiliary stability theorem for a Lyapunov function of a differential inclusion. Then, we apply this theorem to our function $L$ and obtain the stability of Nash equilibria.

First, to prove the Lipschitz continuity of the function $L$ in (6), we should notice that

$$L(\mathbf{x}) = \sum_{p \in \mathcal{P}} \max_{b \in \mathcal{A}^p} L_b^p(\mathbf{x})$$

where the function $L_b^p : \mathcal{X} \to \mathbb{R}$ is given by

$$
\begin{aligned}
L_b^p(\mathbf{x}) &:= \sum_{a \in \mathcal{A}^p} x_a^p \int_0^{F_b^p(\mathbf{x}) - F_a^p(\mathbf{x})} (F_b^p(\mathbf{x}) - F_a^p(\mathbf{x}) - q)Q'(q)dq \\
&= \sum_{a \in \mathcal{A}^p \setminus b} x_a^p \left[ Q(F_b^p(\mathbf{x}) - F_a^p(\mathbf{x}))(F_b^p(\mathbf{x}) - F_a^p(\mathbf{x})) - \int_0^{F_b^p(\mathbf{x}) - F_a^p(\mathbf{x})} qQ'(q)dq \right]
\end{aligned}
$$

for each $p \in \mathcal{P}$ and $b \in \mathcal{A}^p$. This relationship between $L(\mathbf{x})$ and $L_b^p(\mathbf{x})$ is obvious when we interpret the tBRD as a standard BRD with a stochastic status-quo bias. For simplicity, assume a single population and drop the superscript for population. Consider the players who get a revision opportunity under the standard BRD. If they switch to action $b$, then the total net increase of their payoffs is $L_b(\mathbf{x})$. This is maximized if they take the best response action $b \in b(\mathbf{x})$ as it gives the largest payoff; $L(\mathbf{x})$ is this maximized total net increase of payoffs.

Under Assumption 2, $L_b^p(\mathbf{x})$ is Lipschitz continuous in $\mathbf{x} \in \mathcal{X}$ for each action $b \in \mathcal{A}^p$. Thus $L(\mathbf{x})$ is also Lipschitz continuous in $\mathbf{x} \in \mathcal{X}$. Furthermore, $L_b^p(\mathbf{x}_t)$ is Lipschitz continuous in $t \in [0, \infty)$ on a Carathéodory (and thus Lipschitz continuous) solution $\{\mathbf{x}_t\}$. This implies the same Lipschitz continuity of $L(\mathbf{x}_t)$. Furthermore, it follows that

$$\dot{L}(\mathbf{x}_t) = \sum_{p \in \mathcal{P}} \dot{L}_b^p(\mathbf{x}_t) \quad \text{for any } b \in b^p(\mathbf{x}_t) \text{ and almost all } t \in [0, \infty)$$

from a version of Danskin's Envelope Theorem:

**Theorem 9** (Hofbauer and Sandholm, 2009: Theorem A.4). *For each element $z$ in a set $Z$, let $g_z : [0, \infty) \to \mathbb{R}$ be Lipschitz continuous. Let*

$$g_*(t) = \max_{z \in Z} g_z(t) \text{ and } Z_*(t) = \arg\max_{z \in Z} g_z(t).$$

*Then $g_* : [0, \infty) \to \mathbb{R}$ is Lipschitz continuous, and for almost all $t \in [0, \infty)$, we have that $\dot{g}_*(t) = \dot{g}_z(t)$ for each $z \in Z_*(t)$.*

Now based on this fact, we would proceed to prove that our $L$ is a Lyapunov function and guarantees the asymptotic stability of Nash equilibria. But as the tBRD is a differential inclusion and not a differential equation, we need a DI version of the stability theorem of a Lyapunov function:

**Theorem 10.** *Let $A$ be a closed subset of a compact space $\mathcal{X}$ and $A'$ be a neighborhood of $A$. Suppose two continuous functions $W : \mathcal{X} \to \mathbb{R}$ and $\tilde{W} : \mathcal{X} \to \mathbb{R}$ satisfy (i) $W(\mathbf{x}) \geq 0$ and $\tilde{W}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathcal{X}$ and (ii) $W^{-1}(0) = \tilde{W}^{-1}(0) = A$. In addition, assume $W$ is Lipschitz continuous in $\mathbf{x} \in \mathcal{X}$ with Lipschitz constant $K \in (0, \infty)$. If any Carathéodory solution $\{\mathbf{x}^t\}$ starting from $A'$ satisfies*

$$\dot{W}(\mathbf{x}^t) \leq -\tilde{W}(\mathbf{x}^t) \quad \text{for almost all } t \in [0, \infty), \tag{12}$$

*then $A$ is asymptotically stable and $A'$ is its basin of attraction.*[31]

*Proof.* First of all, the fact that $W(\mathbf{x}^t)$ is nonincreasing and the Lyapunov stability of $A$ follows from the fact that $\dot{W}(\mathbf{x}^t) \leq -\tilde{W}(\mathbf{x}^t) \leq 0$ for almost all $t$. The fact that $W(\mathbf{x}^t)$ is nonincreasing and property (i) of $W(\mathbf{x}^t)$ jointly imply its convergence.

Now we want to prove that $W(\mathbf{x}^t)$ can be less than any small positive number after a sufficiently long time $t$. Suppose that there is a positive number $l > 0$ such that $W(\mathbf{x}^t) \geq l$ for all $t \geq 0$ on a Carathéodory solution $\{\mathbf{x}^t\}$.

Lipschitz continuity of $W$ implies

$$d(\mathbf{x}, A) := \min_{\mathbf{y} \in A} |\mathbf{x} - \mathbf{y}| \leq 0.5l/K \quad \Rightarrow \quad W(\mathbf{x}) \leq 0.5l < l.$$

Actually, this assumption implies the existence of $\mathbf{y}$ in the compact set $A$ such that $|\mathbf{x} - \mathbf{y}| \leq 0.5l/K$; then it follows that $W(\mathbf{x}) = |W(\mathbf{x}) - W(\mathbf{y})| \leq K|\mathbf{x} - \mathbf{y}| \leq 0.5l$ from the Lipschitz continuity of $W$ and properties (i,ii) of $W$.

So the Carathéodory solution should satisfy $d(\mathbf{x}^t, A) > 0.5l/K$. Consider a closed set $\check{A}$ defined as

$$\check{A} = \{\mathbf{x} \in \mathcal{X} | d(\mathbf{x}, A) \geq 0.5l/K\}.$$

Then the minimum of $\tilde{W}$ exists in this set $\check{A}$ and

$$\mu := \min_{\mathbf{x} \in \check{A}} \tilde{W}(\mathbf{x}) > 0,$$

since $\check{A}$ is a compact set and the minimizer belongs to $\check{A}$ and thus not to $A = \tilde{W}^{-1}(0)$. As $\mathbf{x}^t \in \check{A}$, we have $-\tilde{W}(\mathbf{x}^t) \leq -\mu$. Hence (12) implies

$$W(\mathbf{x}^t) - W(\mathbf{x}^0) \leq -\int_0^t \tilde{W}(\mathbf{x}^s)ds \leq -\mu t,$$

$$\therefore \quad W(\mathbf{x}^t) \leq W(\mathbf{x}^0) - \mu t$$

for all $t \in [0, \infty)$. As $\mu > 0$, this suggests that $W(\mathbf{x}^t) < 0$ for sufficiently large $t > W(\mathbf{x}^0)/\mu$, contradicting property (i) of $W$. Hence for any positive number $l > 0$, we can find a time $T$ such that we have $W(\mathbf{x}^t) < l$ for all $t \geq T$.

We therefore conclude that any Carathéodory solution $\{\mathbf{x}_t\}$ starting from $A'$ satisfies

$$\lim_{t \to \infty} W(\mathbf{x}_t) = 0,$$

and converges to the set $A = W^{-1}(0)$. $\square$

Now we apply this theorem to prove the global asymptotic stability of Nash equilibria in stable games from a Lyapunov function $L$.

---

[31] This is a modification of Smirnov (2001, Theorem 8.2), where $A$ is a singleton consisting of a rest point of a differential inclusion and an upper Dini derivative is used in place of the assumption that the Lyapunov function $W$ is Lipschitz continuous.

*Proof of Theorem 4.* We show the function $L$ is a strictly decreasing Lyapunov function with $L^{-1}(0) = \text{NE}(\mathbf{F})$. First of all, since the integrand $(\breve{F}_a^p(\mathbf{x}) - q)Q'(q)$ is non-negative on the support of $Q$, the value of $L$ is always non-negative. Besides, since the (first) integral is zero if $F_*^p(\mathbf{x}) = F_a^p(\mathbf{x})$ and positive otherwise, we have $L^{-1}(0) = \text{NE}(\mathbf{F})$.

Consider an arbitrary Carathéodory solution $\{\mathbf{x}^t\}$ starting from a point $\mathbf{x}^0 \in \mathcal{X}$. For almost all time, the solution is differentiable in time and the transition vector satisfies (2) and the time derivative of $L$ equals $\sum_p \dot{L}_{b^p}^p$ for any $b^p \in b^p(\mathbf{x})$. Fix such a moment of time $t$ arbitrarily and henceforth drop the time index $t$. The transition vector $\dot{\mathbf{x}}$ satisfies[32]

$$\dot{\mathbf{x}}^p = \sum_{a \in \mathcal{A}^p} x_a^p Q_a^p (\mathbf{y}_a^p - \mathbf{e}_a^p) \qquad \text{with some } \mathbf{y}_a^p \in B^p(\mathbf{x}) \text{ for each } a \in \mathcal{A}^p.$$

Since $y_{ab}^p > 0$ only if $b \in b^p(\mathbf{x})$, Theorem 9 implies

$$\dot{F}_*^p = \sum_{b \in \mathcal{A}^p} y_{ab}^p \dot{F}_b^p = \sum_{b \in \mathcal{A}^p} y_{ab}^p DF_b^p \dot{\mathbf{x}} = \mathbf{y}_a^p \cdot D\mathbf{F}^p \dot{\mathbf{x}},$$

$$\therefore \frac{d}{dt}\breve{F}_a^p = \dot{F}_*^p - \dot{F}_a^p = (\mathbf{y}_a^p - \mathbf{e}_a^p) \cdot D\mathbf{F}^p \dot{\mathbf{x}}.$$

The time derivative of $L$ at this time $t$ is thus

$$\dot{L} = \sum_{p \in \mathcal{P}} \dot{L}_b^p(\mathbf{x}_t) \quad \text{with any } b \in b^p(\mathbf{x}_t)$$

$$= \sum_{p \in \mathcal{P}} \sum_{a \in \mathcal{A}^p} \dot{x}_a^p \int_0^{\breve{F}_a^p} (\breve{F}_a^p(\mathbf{x}) - q)Q'(q)dq + \sum_{p \in \mathcal{P}} \sum_{a \in \mathcal{A}^p} x_a^p Q_a^{p\prime} \frac{d}{dt}\breve{F}_a^p$$

$$= -\sum_{p \in \mathcal{P}} \sum_{a \in \mathcal{A}^p} x_a^p Q_a^p \int_0^{\breve{F}_a^p} (\breve{F}_a^p - q)Q'(q)dq + \sum_{p \in \mathcal{P}} \sum_{a \in \mathcal{A}^p} x_a^p Q_a^p (\mathbf{y}_a^p - \mathbf{e}_a^p) \cdot D\mathbf{F}^p \dot{\mathbf{x}}$$

$$= -\tilde{L} + \dot{\mathbf{x}} \cdot D\mathbf{F}\dot{\mathbf{x}}$$

$$\leq -\tilde{L},$$

where

$$\tilde{L} := \sum_{p \in \mathcal{P}} \sum_{a \in \mathcal{A}^p} x_a^p Q_a^p \int_0^{\breve{F}_a^p} (\breve{F}_a^p - q)Q'(q)dq.$$

Notice that the last weak inequality comes from the definition of a stable game and $\mathbf{1} \cdot \dot{\mathbf{x}}^p = 0$.

Finally, the function $\tilde{L}$ is always non-negative for the same reason as $L \geq 0$; in particular, it is positive when $\mathbf{x}$ is not a Nash equilibrium, and zero when it is a Nash equilibrium. So the function $L$ is a strict Lyapunov function and satisfies the assumptions in Theorem 10.

We therefore conclude that $L^{-1}(0) = \text{NE}(\mathbf{F})$ is asymptotically stable in the whole state space $\mathcal{X}$. □

## A.4 Theorem 5 (stability of a regular ESS)

**Lemma 1.** *Let $\mathbf{x}^* \in \mathcal{X}$ be a regular ESS. Then there is a neighborhood $O \subset \mathcal{X}$ of $\mathbf{x}^*$ and a constant $C > 0$ such that for each population $p \in \mathcal{P}$*

$$(i) \sum_{b \in U^p} \dot{x}_b^p = -\tilde{Q}^p(\mathbf{x}) \sum_{b \in U^p} \tilde{x}_b^p, \quad (ii) \ \dot{\mathbf{x}}^p \cdot D\mathbf{F}^p \dot{\mathbf{x}}^p \leq C \sum_{b \in U^p} \tilde{x}_b^p$$

---

[32]Here we omit $\mathbf{x}$ from the arguments of functions on $\mathcal{X}$ or $\mathcal{X}^p$, and let $Q_a^p = Q(\breve{F}_a^p(\mathbf{x}))$ and $Q_a^{p\prime} = Q'(\breve{F}_a^p(\mathbf{x}))$.

*Proof.* First, since a regular ESS is a quasi-strict equilibrium, the support of $\mathbf{x}^*$ coincides with the set of the pure best responses $b(\mathbf{x}^*)$; namely, $U^p = \mathcal{A}^p \setminus b^p(\mathbf{x}^*)$. Furthermore, by continuity of $\mathbf{F}^p$, there is a neighborhood $O^p$ of $\mathbf{x}^*$ where any suboptimal action $b \in \mathcal{A}^p \setminus b^p(\mathbf{F}^p) = U^p$ remains suboptimal and $D\mathbf{F}^p$ is negative definite on $T\mathcal{X} \cap \mathbb{R}^A_U$. As the dynamic of any suboptimal action $b$ is $\dot{x}^p_b = -Q(\check{F}^p_b(\mathbf{x}))x^p_b$, we obtain

$$\sum_{b \in U^p} \dot{x}^p_b = -\sum_{b \in U^p} Q(\check{F}^p_b(\mathbf{x}))x^p_b = -\tilde{Q}^p(\mathbf{x}) \sum_{b \in U^p} \tilde{x}^p_b. \tag{13}$$

As the revision rate has an upper bound 1 by Assumption 2, $\tilde{Q}^p(\mathbf{x})$ is at most 1. Then it follows that

$$\sqrt{\sum_{b \in U^p} |\dot{x}^p_b|} \leq \sum_{b \in U^p} |\dot{x}^p_b| = -\sum_{b \in U^p} \dot{x}^p_b = \tilde{Q}^p(\mathbf{x}) \sum_{b \in U^p} \tilde{x}^p_b \leq \sum_{b \in U^p} \tilde{x}^p_b.$$

According to Sandholm (2010a, pp.43-44), this and the local negative definiteness of $D\mathbf{F}^p$ jointly imply the existence of a positive constant $C^p > 0$ such that

$$\dot{\mathbf{x}}^p \cdot D\mathbf{F}^p \dot{\mathbf{x}}^p \leq C^p \sqrt{\sum_{b \in U^p} |\dot{x}^p_b|} \leq C^p \sum_{b \in U^p} \tilde{x}^p_b \tag{14}$$

at any point in the neighborhood $O^p$. Take the intersection of all $O^p$ ($p \in \mathcal{P}$) as $O$ and the largest number of all $C^p$ ($p \in \mathcal{P}$) as $C$. Then (13) and (14) imply equations (i) and (ii). $\qquad \square$

We use this constant $C$ to define the Lyapunov function $L^*$ for the regular ESS $\mathbf{x}^*$ and focus on this neighborhood $O$ as the basin of attraction to $\mathbf{x}^*$.

*Proof of Theorem 5.* From the calculation in the proof of Theorem 4, the time derivative of the original Lyapunov function $L$ is

$$\dot{L} = -\tilde{L} + \sum_{p \in \mathcal{P}} \tilde{Q}^p \dot{\mathbf{x}}^p \cdot D\mathbf{F}^p \dot{\mathbf{x}}^p.$$

Hence we have

$$\dot{L}^* = -\tilde{L} + \sum_{p \in \mathcal{P}} \tilde{Q}^p \dot{\mathbf{x}}^p \cdot D\mathbf{F}^p \dot{\mathbf{x}}^p + C \sum_{p \in \mathcal{P}} \sum_{b \in U^p} \dot{x}^p_b.$$

Lemma 1 implies

$$\dot{L}^* = -\tilde{L} + \sum_{p \in \mathcal{P}} \tilde{Q}^p \left( \dot{\mathbf{x}}^p \cdot D\mathbf{F}^p \dot{\mathbf{x}}^p - C \sum_{b \in U^p} \tilde{x}^p_b \right) \leq -\tilde{L}$$

in the neighborhood $O$ of $\mathbf{x}^*$. Then, Theorem 10 guarantees the local asymptotic stability of $\mathbf{x}^*$. $\qquad \square$

# Acknowledgement

# References

BALKENBORG, D., J. HOFBAUER, AND C. KUZMICS (2012): "Refined Best-Response Correspondence and Dynamics," *Theoretical Economics*, forthcoming.

BENAÏM, M., J. HOFBAUER, AND S. SORIN (2005): "Stochastic Approximations and Differential Inclusions," *SIAM Journal of Control and Optimization*, 44(1), 328–348.

BLUME, L., A. BRANDENBURGER, AND E. DEKEL (1991): "Lexicographic Probabilities and Equilibrium Refinements," *Econometrica*, 59(1), 81–98.

CRESSMAN, R. (2003): *Evolutionary dynamics and extensive form games*. MIT Press.

EREV, I., AND A. E. ROTH (1998): "How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, 88(4), 848–881.

GILBOA, I., AND A. MATSUI (1991): "Social Stability and Equilibrium," *Econometrica*, 59(3), 859–867.

GOLMAN, R., AND S. PAGE (2010): "Basins of attraction and equilibrium selection under different learning rules," *Journal of Evolutionary Economics*, 20, 49–72.

HARTMAN, R. S., M. J. DOANE, AND C.-K. WOO (1991): "Consumer Rationality and the Status Quo," *Quarterly Journal of Economics*, 106(1), 141–162.

HOFBAUER, J. (1995): "Stability for the best response dynamics," mimeo, University of Vienna.

HOFBAUER, J., AND W. H. SANDHOLM (2009): "Stable games and their dynamics," *Journal of Economic Theory*, 144(4), 1665–1693.

HOFBAUER, J., AND K. SIGMUND (1998): *Evolutionary games and population dynamics*. Cambridge University Press.

HONKAPOHJA, S., AND T. ITO (1983): "Stability with Regime Switching," *Journal of Economic Theory*, 29, 22–48.

KLEMPERER, P. (1995): "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *Review of Economic Studies*, 62, 515–539.

KOJIMA, F. (2009): "A Note on Dynamics in Games," mimeo, Stanford University.

KOJIMA, F., AND S. TAKAHASHI (2007): "Anti-Coordination Games And Dynamic Stability," *International Game Theory Review*, 9(4), 667–688.

KUZMICS, C. (2011): "On the Elimination of Dominated Strategies in Stochastic Models of Evolution with Large Populations," *Games and Economic Behavior*, 72(2), 452–66.

LIPMAN, B. L., AND R. WANG (2000): "Switching Costs in Frequently Repeated Games," *Journal of Economic Theory*, 93(2), 149–190.

LIPMAN, B. L., AND R. WANG (2009): "Switching costs in infinitely repeated games," *Games and Economic Behavior*, 66, 292–314.

LOU, Y., Y. YIN, AND S. LAWPHONGPANICH (2010): "Robust congestion pricing under boundedly rational user equilibrium," *Transportation Research Part B*, 44, 15–28.

MADRIAN, B. C., AND D. F. SHEA (2001): "The Power of Suggestion: Inertia in 401 (k) Participation and Savings Behavior," *Quarterly Journal of Economics*, 116(4), 1149–1187.

MASATLIOGLU, Y., AND E. A. OK (2005): "Rational choice with status quo bias," *Journal of Economic Theory*, 121(1), 1–29.

MATSUI, A. (1992): "Best response dynamics and socially stable strategies," *Journal of Economic Theory*, 57(2), 343–362.

NORMAN, T. W. (2009): "Rapid evolution under inertia," *Games and Economic Behavior*, 66(2), 865–879.

——— (2010): "Cycles versus equilibrium in evolutionary games," *Theory and Decision*, 69(2), 167–182.

ORTOLEVA, P. (2010): "Status quo bias, multiple priors and uncertainty aversion," *Games and Economic Behavior*, 69(2), 411–424.

OYAMA, D., W. H. SANDHOLM, AND O. TERCIEUX (2010): "Sampling Best Response Dynamics and Deterministic Equilibrium Selection," mimeo, University of Tokyo, University of Wisconsin-Madison, Paris School of Economics and CNRS.

ROTH, G., AND W. H. SANDHOLM (2012): "Stochastic Approximations with Constant Step Size and Differential Inclusions," *SIAM Journal on Control and Optimization*, forthcoming.

SAGI, J. S. (2006): "Anchored preference relations," *Journal of Economic Theory*, 130(1), 283 – 295.

SAMUELSON, W., AND R. ZECKHAUSER (1988): "Status quo bias in decision making," *Journal of Risk and Uncertainty*, 1(1), 7–59.

SANDHOLM, W. H. (2001): "Potential Games with Continuous Player Sets," *Journal of Economic Theory*, 97(1), 81–108.

——— (2002): "Evolutionary Implementation and Congestion Pricing," *Review of Economic Studies*, 69, 667–689.

——— (2005): "Negative Externalities and Evolutionary Implementation," *Review of Economic Studies*, 72(3), 885–915.

——— (2010a): "Local stability under evolutionary game dynamics," *Theoretical Economics*, 5(1), 27–50.

——— (2010b): *Population games and evolutionary dynamics*. MIT Press.

SANDHOLM, W. H., E. DOKUMACI, AND F. FRANCHETTI (2012): "Dynamo: Diagrams for Evolutionary Game Dynamics," `http://www.ssc.wisc.edu/~whs/dynamo`.

SMIRNOV, G. V. (2001): *Introduction to the Theory of Differential Inclusions*. American Mathematical Society, Providence, RI.

SZETO, W., AND H. K. LO (2006): "Dynamic traffic assignment: properties and extensions," *Transportmetrica*, 2(1), 31–52.

TSAKAS, E., AND M. VOORNEVELD (2009): "The target projection dynamic," *Games and Economic Behavior*, 67(2), 708–719.

VAN DAMME, E. (1991): *Stability and perfection of Nash equilibria*. Springer, 2 edn.

WEIBULL, J. W. (1995): *Evolutionary game theory*. MIT Press, Cambridge, MA.

ZEEMAN, E. (1980): "Population dynamics from game theory," in *Global Theory of Dynamical Systems*, ed. by Z. Nitecki, and C. Robinson, pp. 471–497, Berlin. Springer.

Zusai, D. (2011): "Essays on evolutionary dynamics and applications to implementation problems," Ph.D. thesis, University of Wisconsin-Madison.

———— (2012a): "Aggregate dynamic in a Bayesian game: stochastic status-quo biases and non-aggregability," mimeo, Temple University.

———— (2012b): "Interior convergence under payoff monotone selections and proper equilibrium: application to equilibrium selection," mimeo, Temple University.