Gains in evolutionary dynamics: unifying rational framework for dynamic stability

Dai ZUSAI*

October 13, 2016

Preliminary[1]

Abstract

In this paper, we investigate gains from strategy revisions in deterministic evolutionary dynamics. To clarify the gain from revision, we propose a framework to reconstruct an evolutionary dynamic from optimal decision with stochastic (possibly restricted) available action set and switching cost. Many of major non-imitative dynamics can be constructed in this framework. We formally define net gains from revisions and obtain several general properties of the gain function, which leads to Nash stability of contractive games—generalization of concave potential games—and local asymptotic stability of a regular evolutionary stable state. The unifying framework allows us to apply the Nash stability to mixture of heterogeneous populations, whether heterogeneity is observable or unobservable or whether heterogeneity is in payoffs or in revision protocols. This extends the known positive results on evolutionary implementation of social optimum through Pigouvian pricing to the presence of heterogeneity and non-aggregate payoff perturbations. While the analysis here is confined to general strategic-form games, we finally discuss that the idea of reconstructing evolutionary dynamics from optimization with switching costs and focusing on net revision gains for stability is promising for further applications to more complex situations.

Keywords: evolutionary dynamics, equilibrium stability, Lyapunov function, contractive/negative definite/ stable games, passivity, evolutionary implementation
JEL classification: C73, C62

*Department of Economics, Temple University, 1301 Cecil B. Moore Ave., RA 873 (004-04), Philadelphia, PA 19122, U.S.A. Tel:+1-215-204-8880. E-mail: zusai@temple.edu.

[1] For the most recent version, visit http://sites.temple.edu/zusai/research/gain/.

1 Introduction

Evolutionary dynamics are constructed from aggregation of inertial and myopic decision processes of agents in large population. An agent switches the action only occasionally. Choice of a new action is determined by some "revision protocol," programmed to agents. Revision processes of individuals are aggregated to a differential equation that describes the transition of the action distribution over time. Variety of dynamics comes from variety of revision protocols: best response dynamic from optimization, replicator dynamic from imitation with payoff comparison, etc.

Equilibrium stability is naturally the central issue in deterministic evolutionary dynamics. Like other studies of dynamic systems, the Lyapunov function is the key mathematical tool to establish stability. A Lyapunov function is a scalar-valued function to measure how far the current state is from stable sets. Sandholm (2001) proves that in potential games, the potential function can be used as a Lyapunov function commonly under various major evolutionary dynamics; the potential function is indeed a scalar-valued function to summarize the payoff vector function in the sense that the gradient vector of the potential always coincides with the payoff vector.

But, apart from potential games, a payoff vector function may not be summarized by a scalar-valued function. A game fails to have a potential function just by slight change in the payoff function, unless it keeps "externality symmetry" (Sandholm, 2001).[2] Without a potential function, there is no obvious common choice for a Lyapunov function. Generally the choice of a Lyapunov function is specific to each dynamic. Hofbauer and Sandholm (2009) define a class of games, called *contractive games*, as a generalization of potential games with a concave potential function;[3] a contractive game is characterized by negative marginal effect of a switch of action on the payoff of the new action compared to the old action, which is met by a concave potential game. But a contractive game may not have a potential function. They verify stability of Nash equilibrium set in contractive games for various evolutionary dynamics by finding a Lyapunov function specifically for each canonical dynamic. However, it is left to be unanswered what if those dynamics are mixed, what if there are perturbations to the dynamic, and what is the general economic intuition behind Nash stability.

To obtain a scalar measure of the game dynamic toward equilibrium, it would be natural for economists to look at gains from revisions. If agents have settled down to an equilibrium, there would be no gain from revisions; such a gain would be the incentive for an agent to revise an action, as long as they respond to incentives more or less rationally. Zusai (2015) and Sawa and Zusai (2016) exploit the idea to construct a Lyapunov function in contractive games under versions of best response dynamics.[4] But, to put this idea into other evolutionary dynamics, there seems

[2]It means that the marginal effects of switch from one action to another on payoff difference between arbitrary two actions are completely the same across any pair of actions to switch and any pair of actions whose payoff.

[3]This terminology follows Sandholm (2013). Originally, Hofbauer and Sandholm (2009) call it a stable game. It is also called a negative semidefinite game.

[4]The former introduces stochastic switching cost to the BRD and defines the tempered best response dynamic; it will be discussed as an example of evolutionary dynamics in our framework. The latter considers simultaneous play

to be a puzzling inconsistency between (myopic but) completely rational choice and bounded rationality in those evolutionary dynamics; while a *rational* agent should switch to the optimal action to achieve the greatest net gain from the switch, an agent is actually supposed to take some other suboptimal action under evolutionary dynamics, except the best response dynamic.

To resolve this puzzle and to provide a unified approach for the analysis of equilibrium stability, this paper proposes a new framework to construct evolutionary dynamics. In this paper, we construct a dynamic from optimization protocol like the best response dynamic; but, to justify switches to suboptimal actions, we allow to restrict the available action set and also to impose the switching cost to abandon the current action; they are further allowed to be random.

There were a few studies to put each of these two modifications into some specific dynamic; Zusai (2015) introduces stochastic switching costs into the best response dynamic. Benaïm and Raimond (2010) and Bravo and Faure (2015) consider stochastic restriction on the available action set in fictitious play and reinforcement learning, respectively, though they find that long-run outcomes of these learning processes are captured by the standard best response dynamic. In contrast, this paper combines these two factors and indeed investigate common general properties inherited with all the dynamics that fit into the framework, rather than a particular dynamic. The formulation of evolutionary dynamics from such stochastic *constrained* optimization protocol with switching costs includes many of major deterministic evolutionary dynamics, such as best response dynamic and Smith dynamic.

Under this framework, we can formally define the net gain from revision concretely as the payoff improvement by the revision minus the switching cost. A revising agent optimizes the new action to maximize this net payoff improvement; continuous dependency of the likelihood to actually switch the action is justified by continuous distribution of the stochastic switching cost.

The construction from constrained optimization guarantees stationarity of Nash equilibrium in general and Nash stability in potential games, under mild assumptions: any action should be available with positive probability; between any two actions, the probability that a subset of other actions becomes available should be kept the same, whichever of the two actions has been taken until the revision opportunity; the switching cost must be non-negative but can be any small with some positive probability. These are satisfied by the major dynamics including imitative dynamics such as the replicator dynamic.

We further make another assumption that the probability distribution of available action sets does not change with the social state. Then, our focus is narrowed to non-observatory dynamics such as BRD, pairwise payoff comparison dynamics, etc. While it excludes imitative dynamics, this assumption guarantees that the ordering between actions by expected gains is completely reversed to the ordering by the current payoffs; a player of a worse action enjoys a greater expected gain by switching from this action than a player of a better action. As the constrained optimization protocol dictates agents to switch to (myopically) better actions—though not the best, the total

of multiple games and defines the multitasking BRD by allowing an agent to switch the action only in either one game. This dynamic does not satisfy Assumption A1 in our framework.

expected gain diminishes over time by their switches, unless switches cause positive externality and increase in payoffs refuel the gains. But such positive externality is indeed excluded in a contractive game and the latter possibility is negated by the additional assumption. So, the total expected gain converges to zero over time; then we can use it as a Lyapunov function to verify stability of Nash equilibrium set in contractive games, commonly used for those evolutionary dynamics. Further, with a small modification, the total gain function can be used to verify local stability of a regular ESS.

Our analysis goes step by step; after presenting the construction of evolutionary dynamic from constrained optimization in the next section and verifying general properties of the dynamic itself such as Nash stationarity, we define the expected gain for an individual revising agent and obtain several basic properties of the individual gain function like the reversed ordering in Section 3. Then, we define the aggregate expected gain over the society and then prove Nash stability in contractive games in Section 4.

On this very last step, we find that, apart from specification of a game to be played, evolutionary dynamics in our framework guarantees δ-passivity of the total expected gain function. Fox and Shamma (2013) regard the dynamic interplayed by an evolutionary dynamic and a game as a feedback system where the action distribution is defined as the state of the system and evolves according to a certain aggregate revision process (the evolutionary dynamic) while the payoffs generated from the game influence the state's dynamic by giving the feedback inputs to the state dynamic; in this way, we can separate the evolutionary dynamic as the system that governs the state dynamic and the game as the one that governs the input dynamic. They propose the concept of δ-passivity as a sufficient condition on the state dynamic for stability when the input dynamic satisfies its opposite concept, δ-anti-passivity. With this separation, they extend the Nash stability of contractive games to allow several modifications of the way how agents perceive the payoffs—for example, they respond to weighted moving average of past payoffs, not just the current payoffs; or, they have some predictor of future payoffs and respond to the shock beyond the prediction; these will be briefly reviewed in Section 4. All these modifications keep the δ-anti-passivity and thus retain Nash stability when the game is embedded with δ-passive evolutionary dynamics.

For expository simplicity, all of our theorems are first established for single-population games. In Section 5, we extend it to multi-population games; agents in different populations may have different action sets, different payoff functions and/or different revision protocols. The payoffs may depend on the profile of the action distribution of actions in each population, which means players' heterogeneity is observable when they play the game. Or, it may depend only on the aggregate action distribution in the whole society, when heterogeneity is unobservable. We verify that the overall social dynamic retains δ-passivity if each population dynamic satisfies it. Besides, a game remains to be contractive even if unobservable heterogeneity is introduced to the game, as long as the game is contractive in a single-population homogeneous setting.

Therefore, Nash stability in contractive games is robust to heterogeneity in payoffs and in revision protocols, whether heterogeneity is observable or not. Note that, although replicator dynamic does not meet one of our assumptions and our gain function does not work for this dynamic, an alternative Lyapunov function is found by Hofbauer, Schuster, and Sigmund (1979) and Zeeman (1980).[5] As our extension theorem only depends the existence of a Lyapunov function in a single population setting, it is applicable to the replicator dynamic. So we can allow some agents to follow the replicator dynamic. Sawa and Zusai (2014) verify that any long-run outcomes under an imitative dynamic is robust to heterogeneity in aspiration levels to trigger imitation. Combined this with our result, we can widely extend the range of heterogeneity that guarantees Nash stability of contractive games.

One important application of this extension to the heterogeneous setting is evolutionary implementation of social optimum by Pigouvian pricing. Sandholm (2002, 2005) proposes dynamic Pigouvian pricing to implement the maximum of the total payoffs in the globally asymptotically stable state in evolutionary dynamics. His model allows payoff henterogeneity as an additively separable term in the payoff function, consistent with the discrete choice model in microeconometrics. The Pigovian pricing is calculated only from the common part of the payoff function and independent of the distribution of heterogeneous payoffs or the social dynamic; as long as the common part is precisely identified, the Pigouvian pricing makes the game to possess the total payoff function as a potential function. In particular, if the common payoff function exhibits negative externality like in congestion games, then the total payoff function is concave in the aggregate action distribution; so is the potential function. Therefore, the maximal total payoff is the maximal potential and thus achieved as the stable state thanks to Nash stability of a potential game.

However, when there is perturbation in the actual payoff function from the known common payoff function and it cannot be taken into the calculation of Pigouvian pricing due to misidentification of the payoff function, the game loses the potential function. If negative externality dominates the payoff dependency between heterogeneous agents, then the game is still a contractive game and thus expected to keep Nash stability. Our theorem confirms this promise and extends the applicability of dynamic Pigouvian pricing scheme to the presence of heterogeneity not only in payoffs but also in revision protocols, by providing a unified proof of Nash stability of contractive games. We clarify that the total payoffs can be approximately maximized in the limit state under any evolutionary dynamics in our unified frame, when the payoff perturbation is sufficiently small.

[5] At the time, they did not exactly argue Nash stability in contractive games; but Hofbauer and Sandholm (2009, Sec. 7.2) argue that their Lyapunov function can be used to prove stability of an interior equilibrium in a strict contractive game. See also the bibliographic note on Sandholm (2010b, p.268).

2 Base model

2.1 Single-population game

As a base model, we consider a finite-action game played in the society of continuously many homogeneous agents.[6] The society consists of a unit mass of agents. In this base model, they are homogeneous in the sense that they have the same action set and the same payoff function.

More specifically, each agent chooses an action a from $\mathcal{A} := \{1, \ldots, A\}$. Denote by $x_a \in [0, 1]$ the mass of action-a players in the society. The state of the society is represented by a column vector $\mathbf{x} := (x_1, \ldots, x_A)$ in $\mathcal{X} := \Delta^{\mathcal{A}} = \{\mathbf{x} \in [0, 1]^A \mid \mathbf{1} \cdot \mathbf{x} = 1\}$.[7] The target space of $\mathcal{X}$ is $T\mathcal{X} := \{\mathbf{z} \in \mathbb{R}^A \mid \mathbf{1} \cdot \mathbf{z} = 0\}$.

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

$$\mathbf{F}(\mathbf{x}) := \begin{pmatrix} F_1(\mathbf{x}) \\ \vdots \\ F_A(\mathbf{x}) \end{pmatrix} \quad \text{for each } \mathbf{x} \in \mathcal{X}.$$

Henceforth, we consider a game with a continuously differentiable payoff function.

Assumption F1. The payoff function $\mathbf{F} : \mathcal{X} \to \mathbb{R}^A$ is continuously differentiable.

As usual, a Nash equilibrium is a state where (almost) every agent takes an optimal action. Formally, a social state $\mathbf{x} \in \mathcal{X}$ is a **Nash equilibrium** if $F_a(\mathbf{x}) \geq F_b(\mathbf{x})$ for all $b \in \mathcal{A}$ whenever $x_a > 0$. Denote by $\mathrm{NE}(\mathbf{F})$ the set of Nash equilibria in population game $\mathbf{F}$.

The simplest example of a population game is single-population random matching in a symmetric two-player normal-form game with an $A \times A$ payoff matrix $\mathbf{\Pi}$; the population game is defined by $\mathbf{F}(\mathbf{x}) = \mathbf{\Pi}\mathbf{x}$. $\mathrm{NE}(\mathbf{F})$ coincides with the set of *symmetric* Nash equilibria of $\mathbf{\Pi}$.

Potential games

A population game $\mathbf{F} : \mathcal{X} \to \mathbb{R}^A$ is called a **potential game** if there is a scalar-valued continuously differentiable function $f : \mathcal{X} \to \mathbb{R}$ whose gradient vector always coincides with the relative payoff vector: for all $\mathbf{x} \in \mathcal{X}$, f satisfies[8]

$$\frac{\partial f}{\partial x_a}(\mathbf{x}) = F_a(\mathbf{x}) - \bar{F}(\mathbf{x}) \text{ for all } a \in \mathcal{A}, \qquad \text{i.e., } \left[\frac{df}{d\mathbf{x}}(\mathbf{x})\right]^T = \mathbf{F}(\mathbf{x}) - \bar{F}(\mathbf{x})\mathbf{1}.$$

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

[7] 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 $\mathcal{A} = \{1, \ldots, A\}$, we define $\Delta^{\mathcal{A}}$ as $\Delta^{\mathcal{A}} := \{\mathbf{x} \in [0, 1]^A | \mathbf{1} \cdot \mathbf{z} = A\}$, i.e., the set of all probability distributions on $\mathcal{A}$.

[8] You may notice that $\mathcal{X}$ is only a subspace of $\mathbb{R}^A$. The gradient $\frac{df}{d\mathbf{x}}(\mathbf{x})$ here means the coefficient vector in the linear approximation of change in f on the tangent space of $\mathcal{X}$, i.e., $f(\mathbf{x} + \mathbf{z}) = f(\mathbf{x}) + \frac{df}{d\mathbf{x}}(\mathbf{x})\mathbf{z} + o(|\mathbf{z}|)$ for all $\mathbf{z} \in T\mathcal{X}$.

The class of potential games includes random matching in symmetric games, binary choice games and standard congestion games. The *potential* function f works as a Lyapunov function in a wide range of evolutionary dynamics: replicator, BRD, etc.: see Sandholm (2001).

Contractive games

A population game $\mathbf{F}$ is a **contractive game** if

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

If the strict inequality holds whenever $\mathbf{x}\neq\mathbf{y}$, $\mathbf{F}$ is a strict contractive game.

If $\mathbf{F}$ is C^1, the definition of a contractive game is equivalent to negative semidefiniteness of the Jacobian matrix of $\mathbf{F}$ with respect to the tangent space $T\mathcal{X}$ of the state space $\mathcal{X}$: for any $\mathbf{x}\in\mathcal{X}$,[9]

$$\mathbf{z}\cdot\frac{d\mathbf{F}}{d\mathbf{x}}(\mathbf{x})\mathbf{z}\leq 0 \text{ for any } \mathbf{z}\in T\mathcal{X}.$$

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

2.2 Evolutionary dynamic

General formulation

Here we construct an evolutionary dynamic to formulate transition of the social state over continuous time horizon. In the dynamic, each agent occasionally receives an opportunity to revise the action, following the Poisson process; the arrival rate is fixed at 1.

Upon receipt of a revision opportunity, a revising agent compares payoffs from the current action and from other *available* actions. All the actions are not necessary to be available: the available action set A'_a is a subset of $\mathcal{A}\setminus\{a\}$. Besides, if the revising agent chooses to switch an action, he needs to pay switching cost q. The switching cost is common to all the available actions. The agent chooses the action that yields the greatest net payoff: given payoff vector $\boldsymbol{\pi}\in\mathbb{R}^A$, available action set A'_a and switching cost q, the decision problem of the revising agent who was taking action $a\in\mathcal{A}$ until this revision opportunity is represented as

$$\max\left\{\pi_a, \max_{b\in A'_a}\pi_b - q\right\} \tag{1}$$

Let $\pi_*[A'_a] := \max_{b\in A'_a}\pi_b$, the greatest (gross) payoff among available actions $A'_a\subset\mathcal{A}\setminus\{a\}$, and $b_*(\boldsymbol{\pi};A'_a) := \operatorname{argmax}_{b\in A'_a}\pi_b$ the set of the optimal actions among them, while $\pi_* := \max_{b\in\mathcal{A}}\pi_b$

[9] For differentiable function $\mathbf{f}=(f_i)_{i=1}^m:\mathbb{R}^n\to\mathbb{R}^m$ with input variable denoted by $\mathbf{z}=(z_j)_{j=1}^n$, we denote the Jacobian matrix of $\mathbf{f}$ by $d\mathbf{f}/d\mathbf{z}$; $\partial f_i/\partial z_j$ locates on the i-th row and the j-th column.

the greatest payoff among all the actions $\mathcal{A}$ and $b(\boldsymbol{\pi}) := \max_{b\in\mathcal{A}} \pi_b$ the set of the optimal actions. Note that $\mathbf{x} \in \mathrm{NE}(\mathbf{F})$ is equivalent to $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*[\mathbf{F}(\mathbf{x})])$.[10]

We allow both the available action set A'_a and the switching cost q to be random, though they are determined right after the receipt of the revision opportunity but before choosing a new action. The available action set $A'_a \subset \mathcal{A} \setminus \{a\}$ is drawn from probability distribution $\mathbb{P}_{Aa}$ over $2^{\mathcal{A}\setminus\{a\}}$, the power set of $\mathcal{A} \setminus \{a\}$. The switching cost $q \in \mathbb{R}$ is drawn from probability distribution $\mathbb{P}_Q$ over $\mathbb{R}$. We assume that these distributions are independent of each other, the agent's current action and the social state and also invariant over time.

The revising agent actually switches to either one action in $b_*(\boldsymbol{\pi}; A'_a)$ if $\pi_*[A'_a] - q > \pi_a$; thus the probability of doing so is $\mathbb{P}_Q((0, \pi_*[A'_a] - \pi_a)) =: Q(\pi_*[A'_a] - \pi_a)$.[11] Otherwise, he chooses to maintain the current action a.

Given that a revising agent was taking action a until this revision opportunity, the probability distribution of possible changes in her action upon the revision opportunity takes a form of a probability vector such that

$$\sum_{A'_a \subset \mathcal{A}\setminus\{a\}} \mathbb{P}_A[A'_a] Q(\pi_*[A'_a] - \pi_a))(\mathbf{y}_a[\boldsymbol{\pi}; A'_a] - \mathbf{e}_a) \in T\mathcal{X}. \tag{2}$$

Here $\mathbf{y}_a[\boldsymbol{\pi}; A'_a] \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}; A'_a])$ is a probability vector that has support over $b_*(\boldsymbol{\pi}; A'_a)$. If there are multiple optimum actions that yield the greatest payoff among feasible actions A'_a, this probability vector $\mathbf{y}_a[\boldsymbol{\pi}; A'_a]$ cannot be determined uniquely and thus the transition vector may not be unique. To allow such multiplicity, we formulate the dynamic as a set-valued differential equation, i.e., a differential inclusion. Let $\mathcal{V}_a[\boldsymbol{\pi}]$ be the set of vectors that take the above form:[12]

$$\mathcal{V}_a[\boldsymbol{\pi}] := \sum_{A'_a \subset \mathcal{A}\setminus\{a\}} \mathbb{P}_A[A'_a] Q(\pi_*[A'_a] - \pi_a)(\Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}; A'_a]) - \mathbf{e}_a).$$

Aggregating $\mathcal{V}_a[\boldsymbol{\pi}]$ over the whole population, the transition of the social state is represented by a vector in the set

$$\mathcal{V}(\mathbf{x})[\boldsymbol{\pi}] = \sum_{a\in\mathcal{A}} x_a \mathcal{V}_a(\boldsymbol{\pi}) \subset T\mathcal{X}. \tag{3}$$

Therefore, in population game $\mathbf{F}$, the dynamic of the social state $\mathbf{x}$ is defined as

$$\dot{\mathbf{x}} \in \mathcal{V}(\mathbf{x})[\mathbf{F}(\mathbf{x})] =: \mathcal{V}^{\mathbf{F}}(\mathbf{x}).$$

[10]Consider a $|\mathscr{U}|$-dimensional real space, each of whose coordinate is labeled with either one element of $\mathscr{U} = \{1, \ldots, |\mathscr{U}|\}$. For a set $S \subset \mathscr{U}$, we define a $|S|$-dimensional simplex $\Delta^{|\mathscr{U}|}(S)$ in the $|\mathscr{U}|$-dimensional space as $\Delta^{|\mathscr{U}|}(S) := \left\{\mathbf{x} \in \mathbb{R}_+^{|\mathscr{U}|} \mid \sum_{k\in S} x_k = 1 \text{ and } x_l = 0 \text{ for any } l \in \mathscr{U} \setminus S\right\}$. So it is the set of probability vectors whose support is contained in S.

[11]For expositional simplicity, we assume that the agent always chooses the current action if $\pi_a = \pi_*[A'_a] - q$.

[12]In a vector space $\mathcal{Z}$, with a set $S \subset \mathcal{Z}$, an element $\mathbf{c} \in \mathcal{Z}$ and a scalar $k \in \mathbb{R}$, we define set $k(S + \mathbf{c}) := \{k(\mathbf{z} + \mathbf{c}) \in \mathcal{Z} \mid \mathbf{z} \in S\}$; for sets $S_1, S_2 \subset \mathcal{Z}$, we define set $S_1 + S_2 := \{\mathbf{z}_1 + \mathbf{z}_2 \in \mathcal{Z} \mid \mathbf{z}_1 \in S_1 \text{ and } \mathbf{z}_2 \in S_2\}$.

Assumptions

For our analysis, we make the following assumptions on the revision process. But, as we will see soon, they do not exclude any more major evolutionary dynamics.

Assumption Q1. $Q(0) = \mathbb{P}_Q((-\infty, 0)) = 0$, i.e., $q \geq 0$ almost surely.

Assumption Q2. For any $q > 0$, $Q(q) = \mathbb{P}_Q((-\infty, q)) > 0$.

Assumption A1. For any $a \in \mathcal{A}$ and any $b \in \mathcal{A} \setminus \{a\}$, there is $A'_a \subset \mathcal{A} \setminus \{a\}$ such that $\mathbb{P}_{Aa}[A'_a] > 0$ and $b \in A'_a$.

Assumption A2. For any $a, b \in \mathcal{A}$ and $A_{ab} \subset \mathcal{A} \setminus \{a, b\}$, $\mathbb{P}_{Aa}(\{A'_a \mid A'_a \cap A_{ab} \neq \emptyset\}) = \mathbb{P}_{Ab}(\{A'_b \mid A'_b \cap A_{ab} \neq \emptyset\})$.

The first two assumptions are about the distribution of switching costs. Assumption Q1 excludes the possibility that q adds an additional positive utility to switching from the current action. On the other hand, Assumption Q2 guarantees some positive probability that an agent chooses to switch her action as long as there is an available action strictly better than the current action.

The next two assumptions are about the distribution of available action sets. Assumption A1 implies that, in ex-ante sense before the realization of the available action set, any action will become available with some positive probability. Assumption A2 requires a kind of symmetry: when we compare the distribution of available actions when the current action is a and the one when it is b, they assign the same probability on the event that a certain set of actions become available, as long as the set does not include either a or b.

Our formulation itself imposes a few restrictions. One significant restriction is that the distribution of available action sets does not depend on the social state, which excludes imitative dynamics such as the replicator dynamic or any observatory dynamics in which the availability of actions is based on observation of other agents' choices.

Assumption A3. $\mathbb{P}_{A\cdot}$ is independent of $\mathbf{x}$.

Examples

The construction of evolutionary dynamics from the randomly constrained optimization protocol (1) indeed includes many major dynamics. The following dynamics fit to our construction and satisfy all the assumptions.

Example 1 (Best response dynamics). It is trivial to fit the standard best response dynamic (BRD) $\dot{\mathbf{x}} \in B(\mathbf{x}) - \mathbf{x}$ to our framework just by setting $\mathbb{P}_{Aa}(\mathcal{A}) = 1$ for any $a \in \mathcal{A}$ and $Q(q) = 1$ for all $q > 0$ while $Q(0) = 0$. That is, let every action available almost surely and set switching cost to zero almost surely.

Zusai (2015) generalizes BRD to the tempered BRD (tBRD) by allowing Q to be any increasing function that satisfies Assumptions Q1 and Q2. As the tBRD still assumes $\mathbb{P}_{Aa}(\mathcal{A}) = 1$, our framework could be seen as generalization of the tBRD. But it does indeed expand generality so as to include the following non-optimization based dynamics.

Example 2 (Symmetric pairwise payoff comparison). In a pairwise payoff comparison dynamic, a revising agent compares the current payoff with the payoff from one randomly picked action and switches to the latter action with positive probability (conditional switching rate) if and only if the latter yields a higher payoff than the current payoff. In Smith dynamic, the conditional switching rate is proportional to the payoff difference between the randomly picked action b and the current action a. We can generalize it as the probability is determined by increasing function $Q : \mathbb{R} \to [0,1]$ such as $Q(q) = 0$ for any $q \leq 0$. So, conditional on b randomly picked, the rate of switch from a to b is $Q([\pi_b - \pi_a]_+)$.[13] This dynamic fits with our framework (1) by setting $\mathbb{P}_{Aa}(\{b\}) = 1/(A-1)$ for all $b \in \mathcal{A} \setminus \{a\}$ and defining $\mathbb{P}_Q$ from this switching rate function Q. All the assumptions except Assumption Q2 are guaranteed from this construction. Assumption Q2 is satisfied if $Q(q) > 0$ for all $q > 0$; so does Smith dynamic.

Example 3 (BRD on randomly restricted action set). As already stated, the straight interpretation of our framework on revision protocols is that an agent chooses an action from a randomly restricted action set as to maximize the net payoff. The assumptions on $\mathbb{P}_{A\cdot}$ look very general or vague, while the above examples might not exploit the generality of our framework very well. So it is worth to explore how we can stretch the reach of our framework beyond these known dynamics with concrete examples of $\mathbb{P}_{A\cdot}$.

It is easy to confirm the assumptions on $\mathbb{P}_{A\cdot}$ in the following examples. But, they do not fall neither of the major dynamics mentioned above.

Example 3-i. In a revision opportunity, each action other than the current action becomes available with a certain probability. The availability of each action is independent of those of others and the probability of each action's availability is common to all the actions. This is formulated in $\mathbb{P}_{A\cdot}$ as $\mathbb{P}_{Aa}(A'_a) = p^{\#A'_a}/(1-(1-p)^{A-1})$ with some $p \in (0,1)$. The denominator is to exclude the case that no action becomes available.

Example 3-ii. In a revision opportunity, the number of available actions is determined first: n actions will be available with probability $p_n \in (0,1)$ ($1 \leq n \leq A-1$ and $\sum_{n=1}^{A-1} p_n = 1$). Given n the number of available actions, each possible set of n actions is drawn with equal probability as the available action set for the revising agent. This is formulated as $\mathbb{P}_{Aa}(A'_a) = p_{\#A'_a} / {}_{A-1}C_{\#A'_a}$.

Remark. Excess payoff target dynamics do not fall into this framework as it is. But, the protocol can be interpreted as a constrained optimization protocol by taking the mixed strategy $\mathbf{x}$ (random assignment based on the current action distribution) as the status quo and charging the cost on

[13] The block parentheses with a positive sign $[\cdot]_+$ indicates the non-positive part of the argument: i.e, $[z]_+ = z$ if $z > 0$ and $[z]_+ = 0$ if $z \leq 0$.

switching from this status quo. Then, the gain function satisfies the properties that we will see in Section 4.4.

Nash stationarity in general and stability in potential games

The first three assumptions imply stationarity of the dynamic when all agents are currently playing optimal actions; so, when the dynamic is embedded with a population game, it guarantees stationarity of Nash equilibria. Further, they also imply positive correlation between payoffs of actions and net inflows to them. As Sandholm (2001) shows, the positive correlation immediately implies stability of Nash equilibria in potential games.

Theorem 1 (Best response stationarity)**.** *Assume Assumptions Q1, Q2 and A1. Then, i) state $\mathbf{x} \in \mathcal{X}$ is a stationary state given payoff vector $\boldsymbol{\pi}$, i.e., $\mathcal{V}(\mathbf{x})[\boldsymbol{\pi}] = \{\mathbf{0}\}$, if almost all agents are taking only optimal actions given $\boldsymbol{\pi}$ in state $\mathbf{x}$, i.e., $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$. ii) If $\mathbf{x} \notin \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$, then state $\mathbf{x}$ cannot be stationary, i.e., $\mathbf{0} \notin \mathcal{V}(\mathbf{x})[\pi]$.*

Theorem 2 (Positive correlation)**.** *Assume Assumptions Q1, Q2 and A1. Then, i) for any $\mathbf{x} \in \mathcal{X}, \boldsymbol{\pi} \in \mathbb{R}^A$, we have*

$$\Delta\mathbf{x} \cdot \boldsymbol{\pi} \geq 0 \qquad \text{for any } \Delta\mathbf{x} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}].$$

ii) Moreover, the equality holds if and only if $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$.

Corollary 1 (Nash stationarity; Nash stability in potential game)**.** *Assume that dynamic $\mathcal{V}$ satisfies Assumptions Q1, Q2 and A1 and population game $\mathbf{F}$ satisfies Assumption F1. Then, i) state $\mathbf{x}$ is stationary under the dynamic $\mathcal{V}^{\mathbf{F}}$, i.e., $\mathcal{V}^{\mathbf{F}}(\mathbf{x}) = \{\mathbf{0}\}$ if and only if $\mathbf{x}$ is a Nash equilibrium in $\mathbf{F}$. ii) If $\mathbf{F}$ is a potential game with potential function f, then $\mathrm{NE}(\mathbf{F})$ is globally attracting. Moreover, each local maximizer of f is Lyapunov stable.*

3 Expected gains from revisions in evolutionary dynamics

3.1 Definition of first/second-order expected net gains

If a revising agent chooses to switch an action, the net payoff is $\pi_*[A'_a] - q$; compared to keeping the current action a, the net payoff gain of switching is $\pi_*[A'_a] - \pi_a - q$; the switch is chosen if and only if it is positive. Upon the revision opportunity but before the realization of A'_a and q, the **expected (first-order net) gain** for an action-a player is represented by[14]

$$g_{a*}[\boldsymbol{\pi}] := \sum_{A'_a \subset \mathcal{A} \setminus \{a\}} \mathbb{P}_{Aa}[A'_a] \mathbb{E}_Q[\pi_*[A'_a] - \pi_a - q]_+. \tag{4}$$

This will be the key variable in our analysis. Let $\mathbf{g}_*[\boldsymbol{\pi}] = (g_{a*}[\boldsymbol{\pi}])_{a \in \mathcal{A}}$ be the vector that collects the expected gains of all the actions.

[14] $\mathbb{E}_Q$ is the expected value operator: $\mathbb{E}_Q f(q) = \int_{\mathbb{R}} f(q) d\mathbb{P}_Q(q)$ for integrable function $f : \mathbb{R} \to \mathbb{R}$.

It will be shown that the hypothetical change in gain in the *next* revision opportunity is useful to track the change in the expected gain. We call it the expected second-order gain: specifically, the **expected second-order gain** for an action-a player is defined as

$$h_{a*}[\boldsymbol{\pi}] := \sum_{A'_a \subset \mathcal{A}\setminus\{a\}} P_{Aa}[A'_a] Q(\pi_*[A'_a] - \pi_a)(g_{**}[\boldsymbol{\pi}; A'_a] - g_{a*}[\boldsymbol{\pi}]), \tag{5}$$

where $g_{**}[\boldsymbol{\pi}; A'_a] = \min_{a \in A'_a} g_{a*}[\boldsymbol{\pi}]$. Let $\mathbf{h}_*[\boldsymbol{\pi}] = (h_{a*}[\boldsymbol{\pi}])_{a \in \mathcal{A}}$.

Here, the term $\mathbb{P}_{Aa}[A'_a] Q(\pi_*[A'_a] - \pi_a)$ is the probability that, upon the receipt of the first revision opportunity, they receive available action set A'_a and choose to switch from the current action a to any of the best actions among the available actions. In Theorem 3, we will find that the expected gain has a reversed ordering of the payoff ordering; so, when a revising agent switches to an action that yields the greatest payoff among those available actions, the expected gain becomes the smallest.[15] Hence, $g_{**}[\boldsymbol{\pi}; A'_a]$ is the expected gain of the next switch from a new action chosen in the first revision opportunity to the next action in the second opportunity. So, the difference between $g_{**}[\boldsymbol{\pi}; A'_a]$ and $g_{a*}[\boldsymbol{\pi}]$ measures the difference between the expected gain from possible switches in the first revision opportunity and that in the second opportunity. Thus, $h_{a*}[\boldsymbol{\pi}]$ is the expected change in the expected gains for an agent who was taking action a until the *first* revision opportunity.

3.2 Properties of gain functions

Theorem 3. *Assume Assumptions Q1, Q2, and A1. Then, the expected first-order gain function* $\mathbf{g}_* : \mathbb{R}^A \to \mathbb{R}^A$ *satisfies the following properties.*

g0 *i)* $g_{a*}[\boldsymbol{\pi}] \geq 0$ *for any* $a \in \mathcal{A}$ *and* $\boldsymbol{\pi} \in \mathbb{R}^A$. *ii)* $g_{a*}[\boldsymbol{\pi}] = 0$ *if and only if* $a \in b_*(\boldsymbol{\pi})$.

g1 *Further assume Assumption A2. Consider arbitrary two actions* a, b*; assume* $\pi_a \leq \pi_b$. *Then, i)* $g_{b*}[\boldsymbol{\pi}] \leq g_{a*}[\boldsymbol{\pi}]$. *ii) The inequality is strict if further* $\pi_a < \pi_b$.

g2 *For each* $a \in \mathcal{A}$, *function* $g_{a*} : \mathbb{R}^A \to \mathbb{R}$ *is differentiable almost everywhere in* $\mathbb{R}^A$. *If it is differentiable at* $\boldsymbol{\pi} \in \mathbb{R}^A$,

$$\frac{dg_{a*}}{d\boldsymbol{\pi}}[\boldsymbol{\pi}]\Delta\boldsymbol{\pi} = \sum_{A'_a \subset \mathcal{A}\setminus\{a\}} \mathbb{P}_A[A'_a] Q(\pi_*[A'_a] - \pi_a)(\Delta\pi_*[A'_a] - \Delta\pi_a) \quad \text{for any } \Delta\boldsymbol{\pi} \in \mathbb{R}^A \tag{6}$$

Corollary 2. *Assume that the expected gain function* $\mathbf{g}_* : \mathbb{R}^A \to \mathbb{R}^A$ *satisfies property g1.*[16] *Then, the expected second-order gain function* $\mathbf{h}_* : \mathbb{R}^A \to \mathbb{R}^A$ *satisfies the following property.*

h *i)* $h_{a*}[\boldsymbol{\pi}] \leq 0$ *for any* $a \in \mathcal{A}$ *and* $\boldsymbol{\pi} \in \mathbb{R}^A$. *ii)* $h_{a*}[\boldsymbol{\pi}] = 0$ *if and only if* $a \in b_*(\boldsymbol{\pi})$.

[15] Also, if there are multiple optimum actions, they yield the same expected gain. Thus, the expected gain of the new actions can be determined uniquely.

[16] No other assumptions or properties are needed. Once property g1 is satisfied, any assumption is not needed.

gh *For any* $a \in \mathcal{A}$ *and* $\boldsymbol{\pi} \in \mathbb{R}^A$,

$$h_{a*}[\boldsymbol{\pi}] = \mathbf{z}_a \cdot \mathbf{g}_*[\boldsymbol{\pi}] \qquad \textit{for any } \mathbf{z}_a \in \mathcal{V}_a[\boldsymbol{\pi}].$$

Property g0 guarantees that the expected first-order gain cannot be negative; the revision does not decrease the agent's payoff myopically—as long as the payoff vector is unchanged. Further, the expected gain becomes zero if and only if the current action is indeed myopically optimal among all the actions.

Properties g1 and h are parallel. The former property means that, as long as an agent switches from one action to another, the expected gain cannot increase after the switch. If the new action yields a strictly greater payoff than the old, the expected gain needs to strictly decrease. As a result, the average of the expected gains over all the switchable new actions cannot be greater than the expected gain of switch from the current action; so the expected second-order gain cannot be positive. Further, if the current action is not optimal, then there must be a better action and thus the average of the new expected gains must decrease by the switch; the second-order gain must be negative. And, the opposite is also true.

To understand Property g2, imagine a change in payoff vector by $\Delta\boldsymbol{\pi}$. If a revising agent chooses to switch from current action a to the optimal actions in the realized available action set A'_a, the net gain is $\pi_*[A'_a] - \pi_a - q$; thus, the change in payoff vector $\Delta\boldsymbol{\pi}$ changes this gain by $\Delta\pi_*[A'_a] - \Delta\pi_a$ regardless of q, as long as the switch occurs. Property g2 says that the linear approximation of change in the first-order expected gain $g_{a*}[\boldsymbol{\pi}]$ caused by this change $\Delta\boldsymbol{\pi}$ is obtained by the expected sum of the change in the gain from such a switch $\Delta\pi_*[A'_a] - \Delta\pi_a$ weighted with the current probability of each of possible switches $\mathbb{P}_{Aa}[A'_a]Q(\pi_*[A'_a] - \pi_a)$.

Notice that property g1 means that expected gains makes a reversed ordering of the payoff ordering:

$$\pi_a \geq \pi_b \iff g_{a*}[\boldsymbol{\pi}] \leq g_{b*}[\boldsymbol{\pi}]; \qquad \pi_a = \pi_b \iff g_{a*}[\boldsymbol{\pi}] = g_{b*}[\boldsymbol{\pi}].$$

As we noted in the definition of h_{a*}, when a revising agent optimizes a new action to maximize the payoff among the available actions in A'_a, any of the optimum actions yields the (same) smallest expected gains in A'_a. Therefore, whichever of the (potentially) multiple optimum actions the agent chooses in the switch, the minimum expected gain $g_{**}[\boldsymbol{\pi}; A'_a]$ will be the expected gain from the next switch given that the payoff is unchanged from $\boldsymbol{\pi}$.

4 Aggregate gains and stability of equilibria

4.1 Aggregate gains and δ-passivity

To analyze the overall social dynamic, we aggregate the expected first-order and second-order gains over the whole population. We define the total first/second-order expected gain functions

$G, H : \Delta^{\mathcal{A}} \to \mathbb{R}$ as

$$G(\mathbf{x})[\boldsymbol{\pi}] = \mathbf{x} \cdot \mathbf{g}_*[\boldsymbol{\pi}] \quad \text{and} \quad H(\mathbf{x})[\boldsymbol{\pi}] = \mathbf{x} \cdot \mathbf{h}_*[\boldsymbol{\pi}].$$

Theorem 4. *Consider the aggregates of the expected gain function $\mathbf{g}_* : \Delta^{\mathcal{A}} \times \mathbb{R}^A \to \mathbb{R}^A$ and the expected gain function $\mathbf{h}_* : \Delta^{\mathcal{A}} \times \mathbb{R}^A \to \mathbb{R}^A$ that satisfy properties g0, g2, h and gh.*

Then, the aggregate expected gain function $G : \Delta^{\mathcal{A}} \times \mathbb{R}^A \to \mathbb{R}$ and the aggregate expected second-order gain function $H : \Delta^{\mathcal{A}} \times \mathbb{R}^A \to \mathbb{R}^A$ satisfy the following properties.

G *Property g0 implies i) $G(\mathbf{x}, \boldsymbol{\pi}) \geq 0$ for any $\mathbf{x} \in \Delta^{\mathcal{A}}, \boldsymbol{\pi} \in \mathbb{R}^A$; and, ii) $G(\mathbf{x}, \boldsymbol{\pi}) = 0$ if and only if $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$.*

H *Property h implies i) $H(\mathbf{x}, \boldsymbol{\pi}) \leq 0$ for any $\mathbf{x} \in \Delta^{\mathcal{A}}, \boldsymbol{\pi} \in \mathbb{R}^A$; and, ii) $H(\mathbf{x}, \boldsymbol{\pi}) = 0$ if and only if $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$.*

GH-0 *For any $\mathbf{x} \in \Delta^{\mathcal{A}}, \boldsymbol{\pi} \in \mathbb{R}^A$, property gh implies*

$$H(\mathbf{x})[\boldsymbol{\pi}] = \Delta\mathbf{x} \cdot \mathbf{g}_*[\boldsymbol{\pi}] \qquad \textit{for any } \Delta\mathbf{x} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}].$$

GH-1 *i) For any $\mathbf{x} \in \Delta^{\mathcal{A}}, \boldsymbol{\pi} \in \mathbb{R}^A$, property g2 implies*

$$\frac{\partial G}{\partial \boldsymbol{\pi}}(\mathbf{x}, \boldsymbol{\pi})\Delta\boldsymbol{\pi} = \Delta\mathbf{x} \cdot \Delta\boldsymbol{\pi} \qquad \textit{for any } \Delta\mathbf{x} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}], \Delta\boldsymbol{\pi} \in \mathbb{R}^A.$$

ii) Further assume Assumption A3. Then, properties gh and g2 imply

$$\frac{\partial G}{\partial \mathbf{x}}(\mathbf{x}, \boldsymbol{\pi})\Delta\mathbf{x} = H(\mathbf{x}, \boldsymbol{\pi}) \qquad \textit{for any } \Delta\mathbf{x} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}].$$

It can be easily found that we can retrieve individual gain functions $\mathbf{g}_*$ and $\mathbf{h}_*$ from aggregate ones G and H: as long as $\mathbf{g}_*$ and $\mathbf{h}_*$ are independent of $\mathbf{x}$, we can obtain them from G and H by

$$g_{a*}(\boldsymbol{\pi}) = \mathbf{e}_a \cdot \mathbf{g}_*(\boldsymbol{\pi}) = G(\mathbf{e}_a, \mathbf{x}), \qquad h_{a*}(\boldsymbol{\pi}) = \mathbf{e}_a \cdot \mathbf{h}_*(\boldsymbol{\pi}) = H(\mathbf{e}_a, \mathbf{x}). \tag{7}$$

Here G and H are assumed to be linear in $\mathbf{x}$. The independence or the separability of $\mathbf{x}$ from individual gain functions are implied by Assumption A3. Now one may speculate that properties of $\mathbf{g}_*$ and $\mathbf{h}_*$ in Theorem 3 and Theorem 3 should be now retrieved from properties of G and H in Theorem 4. This speculation is proven to be true, except property g1, as follows.

Theorem 5. *Define the individual gain functions $\mathbf{g}_*, \mathbf{h}_* : \mathbb{R}^A \to \mathbb{R}^A$ from the aggregate gain functions $G, H : \Delta^{\mathcal{A}} \times \mathbb{R}^A \to \mathbb{R}$ by equation 7. Then, they have the following relationships:*

G to g0. *Property G-i implies property g0-i; G-ii implies property g0-ii.*

H to h0. *Property H-i implies property h0-i; G-ii implies property g0-ii.*

GH-0 to gh. *Property GH-0 implies property gh.*

GH-1 to g2. *Property GH-0 implies property g2.*

The corollary below is immediately verified from Theorem 1 and properties G and H in Theorem 4.

Corollary 3. *Suppose that functions G, H satisfy properties G and H. Then, the followings are equivalent to each other: i) $G(\mathbf{x}, \boldsymbol{\pi}) = 0$, ii) $H(\mathbf{x}, \boldsymbol{\pi}) = 0$, iii) $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*(\boldsymbol{\pi}))$, iv) $\mathcal{V}(\mathbf{x})[\boldsymbol{\pi}] = \{\mathbf{0}\}$.*

These properties of G, H imply δ-passivity of the evolutionary dynamic, which is proposed by Fox and Shamma (2013) as the logical source of Nash stability of contractive games:

$$\frac{\partial G}{\partial \mathbf{x}}(\mathbf{x}, \boldsymbol{\pi})\dot{\mathbf{x}} + \frac{\partial G}{\partial \boldsymbol{\pi}}(\mathbf{x}, \boldsymbol{\pi})\Delta\boldsymbol{\pi} \leq \Delta\mathbf{x} \cdot \Delta\boldsymbol{\pi}.$$

With $\Delta\mathbf{x} = \dot{\mathbf{x}}$ and $\dot{\boldsymbol{\pi}} = (d\mathbf{F}/d\mathbf{x})\dot{\mathbf{x}}$, this equation reduces to $\dot{F} \leq \dot{\mathbf{x}} \cdot (d\mathbf{F}/d\mathbf{x})\dot{\mathbf{x}}$. If the game is contractive, this δ-passivity implies that G decreases over time and thus G serves as a Lyapunov function. In the next subsection, we will finalize this idea with a bit stronger version of Lyapunov stability theorem for a differential inclusion.

4.2 Stability of equilibria in contractive games

Now embed an evolutionary dynamic into a population game $\mathbf{F}$. Let $G^{\mathbf{F}}(\mathbf{x}) = G(\mathbf{x}, \mathbf{F}(\mathbf{x}))$ and $H^{\mathbf{F}}(\mathbf{x}) = H(\mathbf{x}, F(\mathbf{x}))$. Then, by the chain rule, we have

$$\frac{d}{dt}G^{\mathbf{F}}(\mathbf{x}_t) = \frac{\partial G}{\partial \mathbf{x}}(\mathbf{x}_t, \mathbf{F}(\mathbf{x}_t))\dot{\mathbf{x}}_t + \frac{\partial G}{\partial \boldsymbol{\pi}}(\mathbf{x}_t, \mathbf{F}(\mathbf{x}_t))\frac{d\mathbf{F}}{d\mathbf{x}}(\mathbf{x}_t)\dot{\mathbf{x}}_t.$$

By property GH-1, this reduces to

$$\frac{d}{dt}G^{\mathbf{F}}(\mathbf{x}_t) = \dot{\mathbf{x}}_t \cdot \frac{d\mathbf{F}}{d\mathbf{x}}(\mathbf{x}_t)\dot{\mathbf{x}}_t + H^{\mathbf{F}}(\mathbf{x}_t). \tag{8}$$

If $\mathbf{F}$ is a contractive game, the negative semidefiniteness of $D\mathbf{F}$ on $T\mathcal{X}$ implies

$$\frac{d}{dt}G^{\mathbf{F}}(\mathbf{x}_t) \leq H^{\mathbf{F}}(\mathbf{x}_t).$$

With other properties of G and H, this implies stability of $\mathrm{NE}(\mathbf{F})$ in a contractive game. Note that, with the Lyapunov stability theorem for a differential inclusion, we allow multiplicity of transition vectors, while utilizing the uniqueness (well definedness) of functions G and H.

Theorem 6. *Consider a contractive game with continuously differentiable payoff function $\mathbf{F}$. Assume an evolutionary dynamic that generate the total expected first/second-order gain functions G, H satisfying properties G, H, and GH-0,1 in Theorem 10. Then, $\mathrm{NE}(\mathbf{F})$ is asymptotically stable under the evolutionary dynamic.*

Proof of Theorem 6. We apply the following version of the Lyapunov stability theorem:

Theorem 7 (Zusai 2015: Theorem 7)**.** *Let A be a closed subset of a compact space $\mathcal{X}$ and A' be a neighborhood of A. Suppose that 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 $\tilde{W}(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \mathcal{X}$ and (ii) $W^{-1}(0) = \tilde{W}^{-1}(0) = A$. In addition, assume that 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*

$$\dot{W}(\mathbf{x}_t) \leq \tilde{W}(\mathbf{x}_t) \quad \text{for almost all } t \in [0, \infty), \tag{9}$$

then A is asymptotically stable and A' is its basin of attraction.

We use function G as the Lyapunov function W in the theorem and function H as the decay rate function $\tilde{W}$. Properties G-i and H-i guarantee condition (i) in the theorem. By properties G-ii and H-ii, condition (ii) is satisfied with the stable set being $A = \mathrm{NE}(\mathbf{F})$. We have verified (9) without making restriction on the initial state $\mathbf{x}_0$; so it holds globally; thus, the basin of attraction is $A' = \mathcal{X}$. □

4.3 Local stability of regular ESS

As Sandholm (2010a) shows, Nash stability of contractive games is a step stone to obtain local stability of evolutionary stable states: the Lyapunov function for a (regular) ESS is obtained by adding one extra term to the one for a contractive game. In the idea presented by Sandholm, the additional term counts the mass of agents who play the actions that are not taken by any positive mass at the ESS. This idea works generally, though it was proven only for each of major evolutionary dynamics: Cressman (1997) for imitative dynamics; Sandholm (2010a) for pairwise comparison dynamics, excess payoff dynamics, and BRD; Zusai (2015) for tBRD.[17] Here we verify the stability of ESS in a general setting.

Similarity between ESS and a contractive game comes from the following analytic definition of an ESS, or precisely so-called a regular ESS: $\mathbf{x}^*$ is a **regular (Taylor) evolutionary stable state** if it satisfies both of the following two conditions:

i) $\mathbf{x}^*$ is a quasi-strict equilibrium:

$$F_b(\mathbf{x}^*) = F_*(\mathbf{x}^*) > F_a(\mathbf{x}^*) \qquad \text{whenever } x_b^* > 0 = x_a^*;$$

ii) $D\mathbf{F}(\mathbf{x}^*)$ is negative definite with respect to $\mathbb{R}^A_{S0}$:

$$\mathbf{z} \cdot D\mathbf{F}(\mathbf{x}^*)\mathbf{z} < 0 \qquad \text{for all } \mathbf{z} \in \mathbb{R}^A_{S0} \setminus \{\mathbf{0}\}.$$

Here S is the set of actions used in $\mathbf{x}^*$, i.e., $S := \{b \in \mathcal{A} \mid x_b^* > 0\}$; and, let

$$\mathbb{R}^A_{S0} := \left\{ \mathbf{z} \in \mathbb{R}^A \mid [z_b = 0 \text{ for any } b \notin S] \text{ and } \sum_{a \in \mathcal{A}} z_a = 0 \right\} = \{\mathbf{z} \in \mathbb{R}^A_0 \mid z_b = 0 \text{ if } x_b^* = 0\}.$$

[17] See also Sawa and Zusai (2016) for the BRD in multitasking setting.

Let $U := \mathcal{A} \setminus S = \{a \in \mathcal{A} \mid x_a^* = 0\}$. Notice that a regular ESS is an isolated Nash equilibrium (Bomze and Weibull, 1995), in the sense that $\mathbf{x}^*$ is the only Nash equilibrium of $\mathbf{F}$ in its small enough neighborhood $X_0 \subset \mathcal{X}$.

Condition ii) suggests that the payoff vector field is contractive around a regular ESS, *if* only the action set is restricted to S^*. With this restriction, the aggregate gain function should decrease strictly and approach to zero over time. However, as other actions than in S^* are indeed allowed, the quadratic term $\dot{\mathbf{x}} \cdot D\mathbf{F}\dot{\mathbf{x}}$ may not be negative semidefinite and thus the aggregate gain might not decrease over time. To overcome this deviation from negative semidefiniteness, Sandholm (2010a) added the mass of players of actions in U. Thanks to continuity of the payoff function and strictness in the first equilibrium condition, these actions in U are strictly worse than the actions in S^*. Thus, the mass of these action players should strictly decrease over time, as long as agents' revisions are consistent with payoffs—we will address it specifically below. By adding a sufficiently large multiple of this term to the aggregate gain function, the decrease in this factor should compensate any potential increase in the quadratic term $\dot{\mathbf{x}} \cdot D\mathbf{F}\dot{\mathbf{x}}$ and keep the whole Lyapunov function decreasing to zero.

Thus, all we need in this argument is the existence of G and H functions and consistency between agents' revisions and payoffs. Like Nash stability of contractive games, we can generalize local stability of ESS to all evolutionary dynamics that fit in our framework.

Theorem 8. *Consider a game with a regular ESS $\mathbf{x}^*$ and an evolutionary dynamic with Assumptions Q1, Q2, and A1. Then, $\mathbf{x}^*$ is (locally) asymptotically stable.*

4.4 Further generalization

Excess payoff target dynamics

In Brown-von Neumann-Nash (BNN) dynamic, a revising agent compares the payoff of each action with the average payoff of the whole population; the payoff difference is called excess payoff. The agent switches to an action with probability proportional to the excess payoff if it is positive. The class of excess payoff target dynamic is generalization of BNN dynamic by allowing the switching probability to be any increasing function of the excess payoff.

Excess payoff target dynamics cannot be fit into our framework, because the benchmark for the revising agent's payoff comparison is the society's average $\mathbf{x} \cdot \boldsymbol{\pi}$, not the payoff of the agent's own current action. This can be interpreted as if the agent who do not choose to "switch" the action would just follow the society's action distribution $\mathbf{x}$ as its own benchmark (mixed) strategy and takes an action according to this mixed strategy. However, with a similar idea as for the other dynamics, we can define the gain function for excess payoff target dynamics.

To better interpret the excess payoff target dynamics, imagine a birth-death process over generations, rather than a revision process of immortal agents. When one agent dies, there is a birth of another new agent; so the mass of agents is fixed constant. The death/birth process follows a Poisson process with arrival rate 1, as the revision opportunity in our standard setting.

The default status-quo strategy for a new born agent is a random choice from the whole population, i.e., the current population state $\mathbf{x}$ as a status-quo mixed strategy. The new born agent can instead take a pure strategy of some particular action, as long as it is available, though it needs to pay the switching cost q. The available action set $A' \subset \mathcal{A}$ is randomly drawn from the distribution $\mathbb{P}_A$ on the power set of $\mathcal{A}$; the switching cost q follows the probability distribution $\mathbb{P}_Q$ on $\mathbb{R}_+$. The new born agent compares the net payoffs from pure strategies of available actions with the expected payoff from the status quo mixed strategy $\mathbf{x}$; then he decides whether to maintain the status quo or to take a costly pure strategy. On $\mathbb{P}_Q$ we keep assuming Assumptions Q1 and Q2. On $\mathbb{P}_A$, while Assumption A2 does not make sense for this situation, we modify Assumption A1:

Assumption (A1′)**.** For any $b \in \mathcal{A}$, there is $A' \subset \mathcal{A}$ such that $\mathbb{P}_A[A'] > 0$ and $b \in A'$.

We assume that, at the moment of decision, the agent does not know which action will be taken if he chooses the status quo. So the (gross) payoff difference between an available pure strategy of action a and the status-quo mixed strategy $\mathbf{x}$ is indeed the relative payoff of the action, i.e., $\hat{\pi}_a := \pi_a - \boldsymbol{\pi} \cdot \mathbf{x}$. In other words, we assume $\mathbf{x}$ as the status quo because the relative payoffs $\hat{\boldsymbol{\pi}}$ is the basis of payoff comparisons in EPT dynamics.

As a result, a revising agent who faces A' as the available action set chooses to take an action from $b_*[A']$ if $\hat{\pi}_*[A']$ is greater than realized switching cost q, i.e., with probability $Q(\hat{\pi}_*[A'])$. Otherwise, the agent picks an action according to $\mathbf{x}$ as a mixed strategy. Thus, the transition of the social state follows

$$\begin{aligned}\dot{\mathbf{x}} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}] &= \sum_{A' \subset \mathcal{A}} \mathbb{P}_A[A'] \left(Q(\hat{\pi}_*[A'])\Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}; A'] + (1 - Q(\hat{\pi}_*[A']))\mathbf{x} \right) - \mathbf{x} \\ &= \sum_{A' \subset \mathcal{A}} \mathbb{P}_A[A'] Q(\hat{\pi}_*[A']) \left(\Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}; A'] - \mathbf{x} \right).\end{aligned}$$

Now we can define an individual's first-order gain $g_*(\boldsymbol{\pi}, \mathbf{x})$ as

$$g_*(\boldsymbol{\pi}, \mathbf{x}) := \sum_{A' \subset \mathcal{A}} \mathbb{P}_A[A'] \mathbb{E}_Q[\hat{\pi}_*[A'] - q]_+.$$

Note that it depends on $\mathbf{x}$ as well as $\boldsymbol{\pi}$, but all the new born agents share the same gain function. As a result, the second order gain is zero; we could say that the choice of action in the first revision does not change the gain in the second revision or simply that mortal agents do not have the second revision.

With the total mass of agents equal to one, the aggregate first-order gain $G(\boldsymbol{\pi}, \mathbf{x})$ is now just equal to the individual first-order gain: $G(\boldsymbol{\pi}, \mathbf{x}) = g_*(\boldsymbol{\pi}, \mathbf{x}) \cdot 1$.

While the aggregate second-order gain cannot be used as a least decay rate function H, we find that the following function H works as the least decay rate function coupled with G as a Lyapunov function to meet properties in Theorem 4.

$$H(\boldsymbol{\pi}, \mathbf{x}) = -\left\{ \sum_{A' \subset \mathcal{A}} \mathbb{P}_A[A'] Q(\hat{\pi}_*[A']) \right\} \left\{ \sum_{A' \subset \mathcal{A}} \mathbb{P}_A[A'] Q(\hat{\pi}_*[A']) \hat{\pi}_*[A'] \right\}.$$

Theorem 9. *For an excess payoff dynamic that satisfy Assumptions Q1, Q2 and A1′, the pair of the above defined functions $G, H : \mathcal{X} \to \mathbb{R}$ satisfy properties G, H, and GH-1 as in Theorem 4.*

(Only) Lyapunov stability in observatory dynamics

Assumption A3 prevents us from including observatory dynamics into our framework: for example, we could define the probability that action b is available $\mathbb{P}_{Aa}(b \in A'_a)$ for imitative dynamic by the probability that a b-player is sampled from the society and then we could define the individual gain function g as in (4) and then the aggregate G. But the essence of such observatory dynamics is that the sampling probability depends on the current actual distribution of actions in the society. Then, Assumption A3 does not hold and indeed the time derivative $\dot{G}$ cannot be clearly separated into the second order gain function H and the second order payoff changes $\dot{\mathbf{x}} \cdot D\mathbf{F}\dot{\mathbf{x}}$ as in (8).

However, the Lyapunov functions have been found for stability of a unique Nash equilibrium, say $\mathbf{x}^*$, under major observatory dynamics; they might be indeed hard to interpret economically.

- Replicator (Taylor and Jonker, 1978): $G(\mathbf{x}; \mathbf{x}^*) = \sum_{a \in S(\mathbf{x}^*)} x_a^* \ln(x_a^*/x_a)$, where $S(\mathbf{x}^*)$ is the set of actions used in $\mathbf{x}^*$.
- Projection dynamic (Nagurney and Zhang, 1997): $G(\mathbf{x}; \mathbf{x}^*) = |\mathbf{x} - \mathbf{x}^*|^2$, where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^A$.

But only Lyapunov stability is obtained; for asymptotic stability, strict contraction is needed. Lahkar and Sandholm (2008) and Sandholm, Dokumacı, and Lahkar (2008) further explore this issue.

Modifications of contractive games

Notice that, even in the cases $\Delta\mathbf{x} \cdot \boldsymbol{\pi} < 0$ is not guaranteed, if we can find a function $f : \mathcal{X} \times \mathbb{R}^A \to \mathbb{R}$ such that i) $f \geq 0$, ii) $f(\mathbf{x}, \boldsymbol{\pi}) = 0$ if and only if $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$ and iii) $\dot{f}(\mathbf{x}_t, \boldsymbol{\pi}_t) \leq -\dot{\mathbf{x}}_t \cdot \dot{\boldsymbol{\pi}}_t$, then we can prove Nash stability by using $G^{\mathbf{F}} + f$ as a Lyapunov function and H as an upper bound function. In this way, Fox and Shamma (2013) verify that Nash stability is maintained by the following modifications to an affine contractive game $\mathbf{F}(\mathbf{x}) = \boldsymbol{\Pi}\mathbf{x} + \mathbf{c}$ where $\boldsymbol{\Pi}$ is a negative definite matrix and $\mathbf{c}$ is a constant payoff vector. See Fox and Shamma (2013, Sec. 4.3.3) for the details.

- Smoothed payoff modification: a revising agent's decision is based on an exponentially weighted moving average of past payoffs. That is, given the payoff stream $\{\boldsymbol{\pi}_t\}_{t \geq 0}$ and discount rate $\lambda > 0$, the perceived payoff vector at time t is

$$e^{-\lambda t}\boldsymbol{\pi}_0 + \int_0^t e^{-\lambda(t-\tau)}\boldsymbol{\pi}_\tau d\tau.$$

- Anticipatory payoff modification: there is a predictor function $\boldsymbol{\pi}^e$ of near-future payoffs such as $\dot{\boldsymbol{\pi}}^e = \lambda(\boldsymbol{\pi} - \boldsymbol{\pi}^e)$. With weight k on the predictor, a revising agent's decision is based on the perceived payoff such as $\boldsymbol{\pi} + k\dot{\boldsymbol{\pi}}^e$. So, if the actual payoff for an action is greater than the predicted payoff, an agent becomes more favor of that action.

5 Extension: heterogeneity and aggregate games

5.1 Multi-population games

The model and all the propositions can be easily extended to a multi-population game in which different populations may have different payoff functions and/or different revision protocols, as long as each population's revision protocol satisfies the assumptions.

To make this concrete, we need to introduce a little more notation. Now we consider a game played by several populations $\mathcal{P} = \{1, \ldots, P\}$. The society consists of a unit mass of the agents, who belong to either one of these P populations; let $m^p > 0$ be the mass of population $p \in \mathcal{P}$ and $\sum_{p\in\mathcal{P}} m^p = 1$. In each population $p \in \mathcal{P}$, all the agents have the same action set $\mathcal{A}^p = \{1, \ldots, A^p\}$; $\mathcal{X}^p := m^p\Delta^{\mathcal{A}^p}$ is the set of all the feasible population states. The social state $\mathbf{x}^{\mathcal{P}} = (\mathbf{x}^p)_{p\in\mathcal{P}} \in \mathbb{R}^{A^{\mathcal{P}}}$ is a collection of each population's state. Here, $A^{\mathcal{P}} = \prod_{p\in\mathcal{P}} A^p$. Denote by $\mathcal{X}^{\mathcal{P}} := \times_{p\in\mathcal{P}} \mathcal{X}^p$ the set of all the feasible social states.

Agents in the same population evaluates the social state based on the same payoff function $\mathbf{F}^p : \mathcal{X}^{\mathcal{P}} \to \mathbb{R}^{A^p}$. So $\mathbf{F}^p(\mathbf{x}^{\mathcal{P}})$ is the payoff vector for agents in population $p \in \mathcal{P}$ when the social state is $\mathbf{x}^{\mathcal{P}}$. We denote by $\boldsymbol{\pi}^{\mathcal{P}}$ the collection of payoff vectors $\boldsymbol{\pi}^p \in \mathbb{R}^{A^p}$ over all the populations $p \in \mathcal{P}$, i.e., $\boldsymbol{\pi}^{\mathcal{P}} = (\boldsymbol{\pi}^p)_{p\in\mathcal{P}} \in \mathbb{R}^{A^{\mathcal{P}}}$.

We construct the dynamic of $\mathbf{x}^{\mathcal{P}}$ as

$$\dot{\mathbf{x}}^{\mathcal{P}} = (\dot{\mathbf{x}}^p)_{p\in\mathcal{P}} \in \mathcal{V}^{\mathcal{P}}(\mathbf{x}^{\mathcal{P}})[\boldsymbol{\pi}^{\mathcal{P}}] := \bigtimes_{p\in\mathcal{P}} \mathcal{V}^p(\mathbf{x}^p)[\boldsymbol{\pi}^p],$$

where $\dot{\mathbf{x}}^p \in \mathcal{V}^p(\mathbf{x}^p)[\boldsymbol{\pi}^p]$ is the transition of population p's state $\mathbf{x}^p$ when the population's current state is $\mathbf{x}^p$ and the payoff vector for the population is $\boldsymbol{\pi}^p$.

We allow different populations to follow different revision protocols, as long as each protocol can be represented by constrained optimization with (possibly) stochastic available set and switching cost. Let $\mathbb{P}^p_Q$ and $\mathbb{P}^p_A$ be the distributions of available sets and of switching costs for population p. As in the single-population base model, we define the expected first and second-order gain functions $\mathbf{g}^p_*, \mathbf{h}^p_* : \mathbb{R}^{A^p} \to \mathbb{R}^{A^p}$ for population p as functions of the population's own payoff vector $\boldsymbol{\pi}^p \in \mathbb{R}^{A^p}$. Similarly, we denote by $G^p, H^P : \mathcal{X}^p \times \mathbb{R}^{A^p} \to \mathbb{R}$ as there aggregates in each population:

$$G^p(\mathbf{x}^p, \boldsymbol{\pi}^p) = \mathbf{x}^p \cdot \mathbf{g}^p_*(\boldsymbol{\pi}^p) \quad \text{and} \quad H^p(\mathbf{x}^p, \boldsymbol{\pi}^p) = \mathbf{x}^p \cdot \mathbf{h}^p_*(\boldsymbol{\pi}^p).$$

Notice that all these four functions are defined from each population's protocol $\mathbb{P}^p_Q, \mathbb{P}^p_A$ and evaluated from the population state and its payoff vector $\mathbf{x}^p, \boldsymbol{\pi}^p$. Thus, as long as each population's protocol satisfies the needed assumptions as specified in the base model, Theorem 3, Corollary 2 and Theorem 4 guarantee that these four functions for the population satisfy all the properties g, h, G, H, GH in these theorems.

Now, to investigate stability in the extended multi-population game on $\mathcal{X}^{\mathcal{P}}$, we further take the sum of all the aggregate gains: define the total aggregate first and second order gain functions $G^{\mathcal{P}}, H^{\mathcal{P}} : \mathcal{X}^{\mathcal{P}} \times \mathbb{R}^{A^{\mathcal{P}}} \to \mathbb{R}$ by

$$G^{\mathcal{P}}(\mathbf{x}^{\mathcal{P}}, \boldsymbol{\pi}^{\mathcal{P}}) := \sum_{p \in \mathcal{P}} G^p(\mathbf{x}^p, \boldsymbol{\pi}^p), \quad \text{and} \quad H^{\mathcal{P}}(\mathbf{x}^{\mathcal{P}}, \boldsymbol{\pi}^{\mathcal{P}}) := \sum_{p \in \mathcal{P}} H^p(\mathbf{x}^p, \boldsymbol{\pi}^p).$$

Theorem 10. *Suppose that, for each population $p \in \mathcal{P}$, the aggregate first and second order gain functions $G^p, H^P : \Delta^{\mathcal{A}^p} \times \mathbb{R}^{A^p} \to \mathbb{R}$ satisfy the properties G, H, and GH-0,1 as in Theorem 4. Then, the total of them $G^{\mathcal{P}}, H^{\mathcal{P}}$ also satisfy the properties G, H, and GH-0,1, respectively.*

Note that, even if a dynamic cannot be fit into our framework of optimization-based revision protocols, we can apply this theorem to the dynamic as long as we can somehow find functions G and H that satisfy these properties for this dynamic. We have seen excess payoff target dynamics for such an example.[18]

5.2 Unobservable heterogeneity in protocols or payoffs

Even though the extension was made straightforward, it has a few particularly interesting implications. Even though different populations are distinguished differently in the above presentation of the model, they need not to be.

More specifically, here we restrict games to aggregate games with additively separable payoffs. In an **aggregate game**, all populations share the same action set $\mathcal{A} = \{1, \ldots, A\}$, i.e., $\mathcal{A}^p = \mathcal{A}$ for all $p \in \mathcal{P}$ and each population's payoff vector depends only on the aggregate action distribution over the society, which is defined as the sum of $\mathbf{x}^1, \ldots, \mathbf{x}^P$, i.e.,

$$\bar{\mathbf{x}} := \sum_{q \in \mathcal{P}} \mathbf{x}^q \in \Delta^{\mathcal{A}},$$

but not on the profile of those in each population $\mathbf{x}^{\mathcal{P}}$. We assume that payoff heterogeneity is additively separable: each population's payoff function $\mathbf{F}^p : \mathcal{X}^{\mathcal{P}} \to \mathbb{R}^A$ consists of common payoff

[18]

function $\mathbf{F}^0 : \Delta^{\mathcal{A}} \to \mathbb{R}^A$ and constant payoff perturbation $\boldsymbol{\theta}^p \in \mathbb{R}^A$:

$$\mathbf{F}^p(\mathbf{x}^{\mathcal{P}}) = \mathbf{F}^0(\sum_{q \in \mathcal{P}} \mathbf{x}^q) + \boldsymbol{\theta}^p.$$

The aggregation represent unobservability of each agent's belonging population when they play the game. In particular, if a game is random matching of a strategic form game with payoff matrix $\boldsymbol{\Pi}$, the common payoff $\boldsymbol{\Pi}\bar{\mathbf{x}}$ is indeed the expected payoff from a match randomly made in the whole society. Note that, like in the last section, different populations may or may not have different revision protocols.

Theorem 11. *If the common payoff function $\mathbf{F}^0 : \Delta^{\mathcal{A}} \to \mathbb{R}^A$ is a contractive game, then the extended payoff function $\mathbf{F}^{\mathcal{P}} = (\mathbf{F}^p)_{p \in \mathcal{P}} : \mathcal{X}^{\mathcal{P}} \to \mathbb{R}^{A^P}$ is also a contractive game.*

Corollary 4. *Introduce additively separable payoff heterogeneity into a contractive game. Even if heterogeneity in payoff and/or revision protocol is not observable when they play the game, Nash stability of the contractive game remains to hold.*

5.3 Robustness of evolutionary implementation

Sandholm's evolutionary implementation: by introducing Pigouvian tax/subsidy into agents' payoffs, the maximum of the total payoff can be maximized in the unique limit state from any initial state, as long as the sum of common payoffs is a concave function of the aggregate action distribution. The global stability does not rely on the distribution of the perturbed payoffs or the specification of the dynamics. The formal proof of the stability relies on Nash stability in a potential game: the Pigouvian pricing makes the game a potential game and then the potential function can be used as the Lyapunov function.

Contractive games are considered as generalization of potential games with concave potential functions. So, it is predicted that small perturbation does not change general stability itself. But it is not formally addressed, because we need to find a Lyapunov function by ourselves once the game becomes not a potential game. But the Lyapunov function depends on specification of the dynamics. Besides, there was no formal study of the case where different agents use different protocols.

In this paper, we have presented how to construct a Lyapunov function in a contractive game even in the presence of heterogeneity in revision protocols and payoff functions. So, now we can formally verify the robustness of global evolutionary implementation of the social optimum through Pigouvian pricing.

Specifically, we consider the following quasi-aggregate game. The society is divided to P populations. Each population's payoff function $\mathbf{F}^p : \mathcal{X}^{\mathcal{P}} \to \mathbb{R}^A$ is

$$\mathbf{F}^p(\mathbf{x}^{\mathcal{P}}) = \mathbf{F}^0(\sum_{q \in \mathcal{P}} \mathbf{x}^q) + \boldsymbol{\theta}^p.$$

Here, payoff heterogeneity enters in the payoff function as an additively separable term.[19] Note that, as in the general model in the last subsection, a different population may follow a different evolutionary dynamic. Define the total common payoff function $\bar{F}^0 : \Delta^{\mathcal{A}} \to \mathbb{R}$ by

$$\bar{F}^0(\bar{\mathbf{x}}) = \bar{\mathbf{x}} \cdot \mathbf{F}^0(\bar{\mathbf{x}}).$$

Correspondingly, we define the total overall payoff function $\bar{F}\mathcal{P} : \mathcal{X}^{\mathcal{P}} \to \mathbb{R}$ as

$$\bar{F}^{\mathcal{P}}(\mathbf{x}^{\mathcal{P}}) = \sum_{q\in\mathcal{P}} \mathbf{x}^q \cdot \mathbf{F}^q(\mathbf{x}^{\mathcal{P}}) = \bar{F}^0(\sum_{q\in\mathcal{P}} \mathbf{x}^q) + \sum_{q\in\mathcal{P}} \mathbf{x}^q \cdot \boldsymbol{\theta}^q.$$

Notice that concavity of $\bar{F}^0$ is equivalent to that of $\bar{F}^{\mathcal{P}}$.

Sandholm (2002) proposes a modification of the game by adding Pigouvian tax/subsidy $\mathbf{T} : \Delta^{\mathcal{A}} \to \mathbb{R}^A$ to the payoff function. For each action $a \in \mathcal{A}$, the Pigouvian tax/subsidy is dynamically adjusted with the change in the aggregate state as

$$T_a(\bar{\mathbf{x}}) := \bar{\mathbf{x}} \cdot \frac{\partial \mathbf{F}^0}{\partial \bar{x}_a}(\bar{\mathbf{x}}) = \sum_{b\in\mathcal{A}} \bar{x}_b \frac{\partial F_b^0}{\partial \bar{x}_a}(\bar{\mathbf{x}}).$$

Calculation of the dynamic Pigouvian pricing only requires identification of the common payoff function $\mathbf{F}^0$ and the precise update of the aggregate state $\bar{\mathbf{x}}$ but does not need knowledge of the underlying dynamic $\mathcal{V}^{\mathcal{P}}$ or the distribution of idiosyncratic payoff vector $(m^p, \boldsymbol{\theta}^p)_{p\in\mathcal{P}}$. With the Pigouvian pricing, the game is modified to $\mathbf{F}^{\mathcal{P}} + \mathbf{T}$; the total overall payoff functions becomes the potential function of the modified game:

$$\begin{aligned}
\frac{\partial \bar{F}^{\mathcal{P}}}{\partial x_a^p}(\mathbf{x}^{\mathcal{P}}) &= F_a^p(\mathbf{x}^{\mathcal{P}}) + \sum_{q\in\mathcal{P}} \mathbf{x}^q \cdot \frac{\partial \mathbf{F}^q}{\partial x_a^p}(\mathbf{x}^p) \\
&= F_a^p(\mathbf{x}^{\mathcal{P}}) + \sum_{q\in\mathcal{P}} \mathbf{x}^q \cdot \frac{\partial \mathbf{F}^0}{\partial \bar{x}_a}(\sum_{q\in\mathcal{P}} \mathbf{x}^q) \\
&= F_a^p(\mathbf{x}^{\mathcal{P}}) + \bar{\mathbf{x}} \cdot \frac{\partial \mathbf{F}^0}{\partial \bar{x}_a}(\sum_{q\in\mathcal{P}} \mathbf{x}^q) = F_a^p(\mathbf{x}^{\mathcal{P}}) + T_a(\sum_{q\in\mathcal{P}} \mathbf{x}^q).
\end{aligned}$$

Therefore, if the total overall payoff function is concave, the maximal total overall payoff is attained in the globally asymptotically stable states under any evolutionary dynamics that guarantee Nash stability in potential games.

One of the most practically important properties of Sandholm's dynamic Pigouvian pricing scheme is no need of specification of the underlying dynamic, thanks to generality of Nash stability in potential games. But it requires accurate knowledge of the common payoff function. It is expected that small perturbation in the payoff function would not lose global stability of Nash

[19] In the situation where payoff heterogeneity is not additively separable and the social planner does not exactly know its distribution, Fujishima (2012) proposes a modified Pigouvian pricing that is combined with estimation of the distribution.

equilibrium. Actually, if the potential function is strictly concave, then the unperturbed game is strictly concave; then, small perturbation should keep strict concavity as we will see, though it would lose a potential function. While Nash stability of concave games has been proven for various major dynamics but not generally, now we verify it in the unified framework to allow heterogeneous mixture of revision protocols. Therefore, we can rigorously extend evolutionary implementability of the social optimum to the heterogeneous setting.

We parameterize the perturbed game $\mathbf{F}^{\mathcal{P}}_\delta = (\mathbf{F}^p_\delta)_{p\in\mathcal{P}} : \mathcal{X}^{\mathcal{P}} \to \mathbb{R}^{AP}$ as follows.

$$\mathbf{F}^p_\delta(\mathbf{x}^{\mathcal{P}}) = \mathbf{F}^0(\sum_{q\in\mathcal{P}} \mathbf{x}^q) + \delta\tilde{\mathbf{F}}^p(\mathbf{x}^{\mathcal{P}}) + \boldsymbol{\theta}^p. \tag{10}$$

Here $\tilde{\mathbf{F}}^p : \mathcal{X}^{\mathcal{P}} \to \mathbb{R}^{A^p}$ represents the perturbation to the common payoff function, which is not observable for the central planner and thus cannot be taken into the Pigouvian pricing. Accordingly, we define the total (perturbed) payoff $\bar{F}^{\mathcal{P}}_\delta : \mathcal{X}^{\mathcal{P}} \to \mathbb{R}$ as

$$\bar{F}^{\mathcal{P}}_\delta(\mathbf{x}^{\mathcal{P}}) = \sum_{p\in\mathcal{P}} \mathbf{x}^p \cdot \mathbf{F}^p_\delta(\mathbf{x}^{\mathcal{P}}).$$

To keep strict contractiveness, we make the following assumptions. The first assumption is a sufficient condition for strict concavity of the total unperturbed common payoff function and for strict contractiveness of the modified game $\mathbf{F}^{\mathcal{P}} + \mathbf{T}$.[20] The second assumption means that the marginal effect of the perturbation on payoffs is bounded.

Assumption F2. i) The common payoff function $\mathbf{F}^0$ is C^1 and satisfies the negative definiteness condition of $d\mathbf{F}^0/d\bar{\mathbf{x}}$: for any $\bar{\mathbf{x}} \in \Delta^{\mathcal{A}}$,[21]

$$\bar{\mathbf{z}} \cdot \frac{d\mathbf{F}^0}{d\bar{\mathbf{x}}}(\bar{\mathbf{x}})\bar{\mathbf{z}} < 0 \text{ for any } \bar{\mathbf{z}} \in T\Delta^{\mathcal{A}}. \tag{11}$$

ii) $\tilde{\mathbf{F}}^{\mathcal{P}}$ is C^1 and there exists an upper bound on $|\partial\tilde{\mathbf{F}}^{\mathcal{P}}/\partial\mathbf{x}^{\mathcal{P}}|$.

Under these assumptions, we can verify that the social optimum can be approximately maximized in the presence of heterogeneity in revision protocols and in unobserved payoffs.

Theorem 12. *Consider the aggregate game as in* (10) *with Assumptions F1 and F2. For any $\varepsilon > 0$ there is $\bar{\delta}(\varepsilon) > 0$ such that*

$$|\delta| < \bar{\delta}(\varepsilon) \quad \Longrightarrow \quad |\lim_{t\to\infty} \bar{F}^{\mathcal{P}}_\delta(\mathbf{x}^{\mathcal{P}}_{\delta,t}) - \max_{\mathbf{x}^{\mathcal{P}}\in\mathcal{X}^{\mathcal{P}}} \bar{F}^{\mathcal{P}}_\delta(\mathbf{x}^{\mathcal{P}})| < \varepsilon,$$

where $\{\mathbf{x}^{\mathcal{P}}_{\delta,t}\}_{t\geq 0}$ is the trajectory starting from an arbitrary initial state $\mathbf{x}^{\mathcal{P}}_{\delta,0} \in \mathcal{X}^{\mathcal{P}}$ in the population game $\mathbf{F}^{\mathcal{P}}_\delta$ under an evolutionary dynamic that satisfies all the assumptions.

[20] The negative definiteness of $d^2\bar{F}^0/d\bar{\mathbf{x}}^2 =$ with respect to $T\Delta^{\mathcal{A}} := \{\mathbf{z} \in \mathbb{R}^A \mid \mathbf{1}\cdot\mathbf{z} = 0\}$ is sufficient for strict concavity of $\bar{F}^0$ in $\Delta^{\mathcal{A}}$. Notice that negative semidefiniteness of the Hessian of a C^2-function is equivalent to concavity of the function.

[21] $T\Delta^{\mathcal{A}}$ is the tangent space of $\Delta^{\mathcal{A}}$: i.e., $T\Delta^{\mathcal{A}} := \{\bar{\mathbf{z}} \in \mathbb{R}^A \mid \mathbf{1}\cdot\bar{\mathbf{z}} = 0\}$.

6 Concluding remarks

In this paper, we propose the notion of gains from revisions in deterministic evolutionary dynamics. To rigorously define it quantitatively, we reconstruct a dynamic from an optimization-based revision protocol. A variety of evolutionary dynamics is retained by allowing stochastic restriction to available actions and also random switching costs. The aggregate net gain serves as a Lyapunov function to prove Nash stability of contractive games; with a small modification, this further extends to local stability of a regular ESS.

Our general approach not only provides general proofs of these fundamental stability theorems with consistent and economic intuitive logic but also extends the results to the heterogeneous setting where different agents follow different revision protocols or their payoff functions differ in the additive separable manner; the heterogeneity may or may not be observable.

While our analysis is confined to a strategic-form game, the author believes that the idea of net gains should be generally applicable to prove global stability of the equilibrium set in a version of contractive games or local stability of a sort of an ESS. For example, Sawa and Zusai (2016) consider the BRD in simultaneous play of multiple games (multitasking BRD); though the mBRD does not fall exactly into our framework presented in this paper to construct evolutionary dynamics, they apply the idea of the net gain to prove Nash stability of contractive games and local stability of a regular ESS.

Perhaps, typically in sequential-move games or repeated games, one might define a version of an ESS that is tailored to each specific situation in order to refine the equilibrium concept.[22] The refinement *à la* ESS may be based on a rough crude idea on how an agent learns plays of others in the society and adjusts the behavior to it, and thus one may hope to justify it by arguing dynamic stability. This paper suggests to formally define such a dynamic from the rough idea and then track down the transition of the net gain from revisions under the dynamic; a proper formulation of the dynamic and the gain concept should yield dynamic stability of the tailor-made equilibrium concept.

A Proofs

A.1 Proof of Theorem 1

Proof. First of all, if $a \in b_*[\boldsymbol{\pi}]$, then $\pi_a = \pi_* \geq \pi_*[A'_a]$ for any $A'_a \subset \mathcal{A} \setminus \{a\}$. Thus, for any of such A'_a, Assumption Q1 implies $Q(\pi_*[A'_a] - \pi_a) = 0$: no switch occurs from a. Hence, we have

$$a \in b_*[\boldsymbol{\pi}] \quad \Longrightarrow \quad \mathcal{V}_a[\boldsymbol{\pi}] = \{\mathbf{0}\}.$$

[22]For example, Fujiwara-Greve and Okuno-Fujiwara (2009) consider a repeated game in which players can choose whether to keep the match or to seek for a new opponent. They define a version of ESS for random matching in large population that play this kind of repeated games: see Fujiwara-Greve and Okuno-Fujiwara (2016).

In particular, this further implies that there is no outflow from the mass of action a-players; since $z_{ba} \geq 0$ always holds in any $\mathbf{z} \in \mathcal{V}_b[\boldsymbol{\pi}]$ as long as $b \neq a$, the assumption $a \in b_*[\boldsymbol{\pi}]$ implies

$$z_a = x_a z_{aa} + \sum_{b \in \mathcal{A} \setminus \{a\}} x_b z_{ba} \geq 0$$

in any $\mathbf{z} \in \mathcal{V}[\boldsymbol{\pi}]$; that is, x_a cannot decrease.

In contrast, consider a suboptimal action $a \notin b_*[\boldsymbol{\pi}]$. Then, Assumption A1 implies the existence of $A'_a \subset \mathcal{A} \setminus \{a\}$ such that $\mathbb{P}_{Aa}(A'_a) > 0$ and $A'_a \cap b_*[\boldsymbol{\pi}] \neq \emptyset$. Then, as $\pi_* = \pi_*[A'_a] > \pi_a$, Assumption Q2 implies $Q(\pi_*[A'_a] - \pi_a) > 0$. Besides, as $\mathbf{y}_a[\boldsymbol{\pi}; A'_a] \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}; A'_a])$ and $b_*[\boldsymbol{\pi}] \supset b_*[\boldsymbol{\pi}; A'_a]$, there exists an action $b \in b_*[\boldsymbol{\pi}]$ such that $y_{ab}[\boldsymbol{\pi}; A'_a] > 0$. Hence, the switch from a to b occurs with positive probability $z_{ab} \geq P_{Aa}(A'_a) Q(\pi_*[\boldsymbol{\pi}; A'_a] - \pi_a) y_{ab}[\boldsymbol{\pi}; A'_a] > 0$, conditional on a being played before the revision opportunity. Notice that, as $b \in b_*[\boldsymbol{\pi}]$, there is no outflow from b as we argued above.

Now, assume $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$; if $x_a > 0$, then $a \in b_*[\boldsymbol{\pi}]$ and thus $\mathcal{V}_a[\boldsymbol{\pi}] = \{\mathbf{0}\}$ as we verified here first. Hence, we have $\mathcal{V}(\mathbf{x})[\boldsymbol{\pi}] = \sum_a x_a \mathcal{V}_a[\boldsymbol{\pi}] = \{\mathbf{0}\}$.

For the converse, assume $\mathbf{x} \notin \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$; then there exists an suboptimal action $a \notin b_*[\boldsymbol{\pi}]$ played by a positive mass of agents $x_a > 0$. As we argued above, for any $\mathbf{y}_a[\boldsymbol{\pi}; A'_a] \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}; A'_a])$ there must be an optimal action $b \in b_*[\boldsymbol{\pi}]$ such that $z_{ab} > 0$. Then, by $x_a > 0$, we have

$$z_b \geq x_b z_{bb} + x_a z_{ab} > 0.$$

So $\mathbf{0} \notin \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}]$. □

A.2 Proof of Theorem 2

Proof. Notice that, for any $\mathbf{z}_a \in \mathcal{V}_a[\boldsymbol{\pi}]$, we have[23]

$$\mathbf{z}_a \cdot \boldsymbol{\pi} = \sum_{A'_a \subset \mathcal{A} \setminus \{a\}} \mathbb{P}_{Aa}[A'_a] Q(\pi_*[A'_a] - \pi_a)(\pi_*[A'_a] - \pi_a).$$

i) Assumption Q1 means that $Q(\pi_*[A'_a] - \pi_a) = 0$ if $\pi_*[\boldsymbol{\pi}; A'_a] < \pi_a$. Hence, the product of these two terms is always non-negative. As $P_{Aa}[\cdot] \geq 0$, this implies $\mathbf{z}_a \cdot \boldsymbol{\pi} \geq 0$. As $\mathcal{V}(\mathbf{x})[\boldsymbol{\pi}] = \sum x_a \mathcal{V}_a[\boldsymbol{\pi}]$, this further implies $\mathbf{z} \cdot \boldsymbol{\pi} = \sum x_a \mathbf{z}_a \cdot \boldsymbol{\pi} \geq 0$ for any $\Delta\mathbf{x} = \sum x_a \mathbf{z}_a \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}]$: part i) is verified.

ii) The "if" part is immediate from Theorem 1-i), as $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$ implies $\mathcal{V}(\mathbf{x})[\boldsymbol{\pi}] = \{\mathbf{0}\}$. For the "only-if" part, assume $\mathbf{x} \notin \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$. Then, there is a suboptimal action $a \notin b_*[\boldsymbol{\pi}]$ played by a positive mass $x_a > 0$. Suboptimality means $\pi_* > \pi_a$. By Assumption A1, there exists an available set A'_a such that $A'_a \cap b_*[\boldsymbol{\pi}] \neq \emptyset$ and $P_{Aa}[A'_a] > 0$; the former implies $\pi_*[A'_a] = \pi_* > \pi_a$. It follows that $Q(\pi_*[A'_a] - \pi_a) > 0$ by Assumption Q2. Therefore, for this A'_a, we have

[23] In comparison with the expected net gain g_{a*} that we will define in the next section, we could call $\mathbf{z}_a \cdot \boldsymbol{\pi}$ the expected gross gain of revision from action a. The only difference is in switching costs.

$P_{Aa}[A'_a]Q(\pi_*[A'_a] - \pi_a)(\pi_*[A'_a] - \pi_a) > 0$; thus, we have $\mathbf{z}_a \cdot \boldsymbol{\pi} > 0$ for this suboptimal action a, regardless of $\mathbf{z}_a \in \mathcal{V}_a[\boldsymbol{\pi}]$. As $x_a > 0$, we have $\mathbf{z} \cdot \boldsymbol{\pi} \geq x_a \mathbf{z}_a \cdot \boldsymbol{\pi} > 0$ for any $\Delta\mathbf{x} = \sum x_a \mathbf{z}_a \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}]$: the contrapositive of the "only-if" part is verified. □

A.3 Proofs of Theorem 3 and Corollary 2

Proof of Theorem 3

Proof. **g0.** i) is immediate from the definition of g_{a*}. Note that the term $\mathbb{E}_Q[\pi_*[A'_a] - \pi_a - q]_+$ reflects the fact that a revising agent switches to a different new action only if $\pi_*[A'_a] \geq \pi_a$; so does this non-negativity of g_{a*}.

For the only-if part of ii), first assume $a \in b_*[\boldsymbol{\pi}]$. Then, $\pi_a \geq \pi_*[A'_a] \geq \pi_*[A'_a] - q$ for any $A'_a \subset \mathcal{A} \setminus \{a\}$, as long as $q \geq 0$. Thus, Assumption Q1 implies $[\pi_*[A'_a] - \pi_a - q]_+ = 0$ almost surely. Thus, $g_{a*}(\boldsymbol{\pi}) = 0$.

For the "if" part of ii), we prove its contrapositive by assuming the existence of $b \in \mathcal{A}$ such that $\pi_b > \pi_a$ and then deriving $g_{a*}(\boldsymbol{\pi}) > 0$. Assumption A1 implies the existence of $A'_a \subset \mathcal{A} \setminus \{a\}$ such that $\mathbb{P}_{Aa}[A'_a] > 0$ such that $b \in A'_a$. In such a set A'_a, we have $\pi_*[A'_a] \geq \pi_b > \pi_a$. Then, by Assumption Q2, it is guaranteed that $Q(\pi_*[A'_a] - \pi_a) > 0$ and thus $\mathbb{E}_Q[\pi_*[A'_a] - \pi_a - q]_+ > 0$. Therefore, we have $g_{a*}(\boldsymbol{\pi}) > 0$.

g1. Make a partition of $\mathcal{A}$ according to $\boldsymbol{\pi}$, say $\mathcal{A}_1, \mathcal{A}_2, \ldots$ such as

$$[a, a' \in \mathcal{A}_i \;\Leftrightarrow \pi_a = \pi_{a'}] \quad \text{and} \quad [a \in \mathcal{A}_i \text{ and } a' \in \mathcal{A}_{i'} \text{ with } i < i' \;\Leftrightarrow\; \pi_a > \pi_{a'}].$$

Let π_i be the payoff obtained from actions in set $\mathcal{A}_i$: i.e., $\pi_i := \pi_a$ with some (and indeed all) $a \in \mathcal{A}_i$. $\mathcal{A}_i$ is the set of actions that yield the i-th greatest payoff π_i among $\{\pi_a \mid a \in \mathcal{A}\}$.

Define $p^i_{A,a}$ by

$$p^i_{A,a} = \mathbb{P}_{A,a}(\{A^i_a \subset \mathcal{A} \setminus \{a\} \mid A^i_a \cup \mathcal{A}_i \neq \emptyset \text{ and } A^i_a \cup \mathcal{A}_j = \emptyset \text{ for any } j < i\}).$$

Given the current action a, the maximal feasible payoff after the revision becomes π_i with probability $p^i_{A,a}$. For set A^i_a, we have $\pi_*[A^i_a] = \pi_i$. Assumption A1 implies $p^1_{A,a} > 0$.

Let $I(a)$ be the index of the partition that a belongs to: i.e., $a \in \mathcal{A}_{I(a)}$. Notice that $g_{a*}(\boldsymbol{\pi}) = \sum_{i=1}^{I(a)-1} p^i_{A,a} \mathbb{E}_Q[\pi_i - \pi_a - q]_+$.

If $\pi_a \leq \pi_b$, then $I(a) \geq I(b)$ as well as

$$\mathbb{E}_Q[\pi_i - \pi_a - q]_+ \geq \mathbb{E}_Q[\pi_i - \pi_b - q]_+. \tag{12}$$

By applying Assumption A2 to $A_{ab} = \bigcup_{i=1}^{I} \mathcal{A}_i$, we obtain $\sum_{i=1}^{I} p^i_{A,a} = \sum_{i=1}^{I} p^i_{A,b}$ for all $I < I(b)$; notice $a, b \notin \bigcup_{i=1}^{I} \mathcal{A}_i$ as $I < I(b) \leq I(a)$. Hence, we have

$$p^i_{A,a} = p^i_{A,b} \quad \text{for all } i < I(b). \tag{13}$$

Combining this with (12), we obtain part i):

$$
\begin{aligned}
g_{a*}(\boldsymbol{\pi}) &= \sum_{i=1}^{I(b)-1} p^i_{A,a}\mathbb{E}_Q[\pi_i - \pi_a - q]_+ + \sum_{i=I(b)}^{I(a)-1} p^i_{A,a}\mathbb{E}_Q[\pi_i - \pi_a - q]_+ \\
\geq g_{b*}(\boldsymbol{\pi}) &= \sum_{i=1}^{I(b)-1} p^i_{A,b}\mathbb{E}_Q[\pi_i - \pi_b - q]_+.
\end{aligned}
$$

Further, if $\pi_a < \pi_b$, then $I(a) > I(b)$. As $\pi_a < \pi_b \leq \pi_1$, Assumption Q2 implies $Q(\pi_1 - \pi_a) > 0$ and thus $\mathbb{E}_Q[\pi_1 - \pi_a - q]_+ > \mathbb{E}_Q[\pi_1 - \pi_b - q]_+$. With the fact $p^1_{A,a} > 0$, it guarantees $g_{a*}(\boldsymbol{\pi}) > g_{b*}(\boldsymbol{\pi})$, part ii) of property g1.

g2. First of all, let $\mathfrak{b}(A'_a) \in A'_a$ be a choice for each non-empty available action set $A'_a \neq \emptyset$; we could call $\mathfrak{b} : 2^{\mathcal{A}\setminus\{a\}} \setminus \emptyset \to \mathcal{A} \setminus \{a\}$ a policy function for a revising agent to make a choice after observing (non-empty) available action set. Let $\mathfrak{B}$ be the set of all such policy functions; notice that it is a finite set as long as $\mathcal{A}$ is a finite set. Define the expected gain from policy $\mathfrak{b}$, $g_{a\mathfrak{b}}$, by

$$
g_{a\mathfrak{b}}(\boldsymbol{\pi}) := \sum_{A'_a \subset \mathcal{A}\setminus\{a\}} P_{Aa}(A'_a)\mathbb{E}_Q[\pi_{\mathfrak{b}(A'_a)} - \pi_a - q]_+.
$$

Under the randomly constrained optimization protocol (1), an agent is supposed to choose $\mathfrak{b}(A'_a)$ from $b_*(A'_a)$; thus,

$$
g_{a*}(\boldsymbol{\pi}) = \max_{\mathfrak{b}\in\mathfrak{B}} g_{a\mathfrak{b}}(\boldsymbol{\pi}).
$$

Specifically, the maximum is attained by a selection $\mathfrak{b}_* \in \mathfrak{B}$ such that $\mathfrak{b}_*(A'_a) \in b_*[A'_a] \subset A'_a$.

With $\mathfrak{b}$ fixed arbitrarily in $\mathfrak{B}$, function $g_{a\mathfrak{b}}$ is differentiable and the derivative is bounded:

$$
\frac{\partial g_{a\mathfrak{b}}}{\partial \pi_b}(\boldsymbol{\pi}) = \begin{cases} -\sum_{A'_a\subset\mathcal{A}\setminus\{a\}} \mathbb{P}_{Aa}[A'_a]Q(\pi_{\mathfrak{b}(A'_a)} - \pi_a) \in [-1, 0] & \text{if } b = a, \\ \mathbb{P}_{Aa}[A'_a]Q(\pi_{\mathfrak{b}(A'_a)} - \pi_a) \in [0, 1] & \text{if } b = \mathfrak{b}(A'_a) \text{ with some } A'_a \subset \mathcal{A} \\ 0 & \text{otherwise.} \end{cases}
$$

$$
\frac{\partial g_{a\mathfrak{b}}}{\partial \boldsymbol{\pi}}(\boldsymbol{\pi})\Delta\boldsymbol{\pi} = \sum_{A'_a\subset\mathcal{A}\setminus\{a\}} \mathbb{P}_{Aa}(A'_a)Q(\pi_{\mathfrak{b}(A'_a)} - \pi_a)(\Delta\pi_{\mathfrak{b}(A'_a)} - \Delta\pi_a).
$$

As the derivative vector of $g_{a\mathfrak{b}}$ is bounded in $[-1, 1]^A$, function $g_{a\mathfrak{b}}$ is Lipschitz continuous on $\mathbb{R}^A$; it implies that the maximal value function g_{a*} is also Lipschitz continuous and differentiable almost everywhere in $\mathbb{R}^A$. The derivative formula (6) is obtained by applying the above equation to $\mathfrak{b} = \mathfrak{b}_*$ and noticing $\pi_{\mathfrak{b}_*(A'_a)} = \pi_*(A'_a)$. □

Proof of Corollary 2

Proof. **h.** Part i) is immediate from g1-i). For part ii, first consider an action a such that $a \in b_*(\boldsymbol{\pi})$. Then, any other action $b \in \mathcal{A}$ yields $\pi_b \leq \pi_a$; by g1-i) again, it is equivalent to $g_{b*}[\boldsymbol{\pi}] \geq g_{a*}[\boldsymbol{\pi}]$; thus, $g_{**}[\boldsymbol{\pi}; A'_a] = \min_{b\in A'_a} g_{b*}[\boldsymbol{\pi}] \geq g_{a*}[\boldsymbol{\pi}]$. So $h_{a*}(\boldsymbol{\pi})$ cannot be strictly negative. With part i) of this property h, we have $h_{a*}[\boldsymbol{\pi}] = 0$. For the opposite, assume $h_{a*}[\boldsymbol{\pi}] < 0$. Then, there must exist A'_a such that $\mathbb{P}_{Aa}[A'_a] > 0, Q(\pi_*[A'_a] - \pi_a) > 0$ and $g_{**}[\boldsymbol{\pi}; A'_a] - g_{a*}[\boldsymbol{\pi}] < 0$. By property g1-ii), it must be the case that $\pi_*[A'_a] > \pi_a$; there exists an action in $A'_a \subset \mathcal{A} \setminus \{a\}$ that is better than a. Therefore, $a \notin b_*[\boldsymbol{\pi}]$.

gh. First of all, any actions in $b_*[\boldsymbol{\pi}; A'_a]$ yield the greatest payoff $\pi_*[A'_a]$ among all the available actions in A'_a; property g1 implies that they yields the smallest first-order gain $g_{**}[\boldsymbol{\pi}; A'_a]$ among them. Thus, any mixture $\mathbf{y}_a[\boldsymbol{\pi}; A'_a] \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}; A'_a)$ satisfy

$$\mathbf{y}_a[\boldsymbol{\pi}; A'_a] \cdot \mathbf{g}_*[\boldsymbol{\pi}] = g_{**}[\boldsymbol{\pi}; A'_a].$$

With the fact $\mathbf{e}_a \cdot \mathbf{g}_*[\boldsymbol{\pi}] = g_{a*}[\boldsymbol{\pi}]$ and the assumption $\mathbf{z}_a \in \mathcal{V}_a[\boldsymbol{\pi}]$, this implies the equation in property gh. □

A.4 Proof of Theorem 4

Proof. **G.** i) is immediately obtained from property g0-i and the fact $\mathbf{x} \in \mathbb{R}^{\mathcal{A}}_+$. For ii), notice that $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$ is equivalent to $x_a > 0 \;\Rightarrow\; a \in b_*[\boldsymbol{\pi}]$; and, by g0-i), $G(\mathbf{x}, \pi) = 0$ is equivalent to $x_a > 0 \;\Rightarrow\; g_{a*}(\boldsymbol{\pi}) = 0$. The equivalence between $a \in b_*[\boldsymbol{\pi}]$ and $g_{a*}(\boldsymbol{\pi}) = 0$ has been established in g0-ii). Thus, combining these, we obtain G-ii).

H. Each of the two parts can be readily verified from property h, similarly to the proof of property G.

GH-0. Since $\mathbf{z} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}] = \sum_{a\in\mathcal{A}} x_a \mathcal{V}_a[\boldsymbol{\pi}]$, there exists $(\mathbf{z}_a)_{a\in\mathcal{A}}$ such that $\mathbf{z} = \sum_{a\in\mathcal{A}} x_a \mathbf{z}_a$ and $\mathbf{z}_a \in \mathcal{V}_a[\boldsymbol{\pi}]$ for each a. By property gh, we have

$$\mathbf{z} \cdot \mathbf{g}_*[\boldsymbol{\pi}] = \sum_{a\in\mathcal{A}} x_a \mathbf{z}_a \cdot \mathbf{g}_*[\boldsymbol{\pi}] = \sum_{a\in\mathcal{A}} x_a h_{a*}[\boldsymbol{\pi}] = H(\mathbf{x})[\boldsymbol{\pi}].$$

GH-1. i) As $G(\mathbf{x}, \boldsymbol{\pi}) = \mathbf{x} \cdot \mathbf{g}_*(\boldsymbol{\pi})$, property g2 implies

$$\begin{aligned}\frac{\partial G}{\partial \boldsymbol{\pi}}(\mathbf{x}, \boldsymbol{\pi})\Delta\boldsymbol{\pi} &= \sum_{a\in\mathcal{A}} x_a \frac{\partial \mathbf{g}_{a*}}{\partial \boldsymbol{\pi}}(\boldsymbol{\pi})\Delta\boldsymbol{\pi} \\ &= \sum_{a\in\mathcal{A}} x_a \sum_{A'_a \subset \mathcal{A}\setminus\{a\}} \mathbb{P}_A[A'_a] Q(\pi_*[A'_a] - \pi_a)(\Delta\pi_*[A'_a] - \Delta\pi_a)\end{aligned}$$

On the other hand, any $\Delta\mathbf{x} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}]$ satisfies

$$\Delta\mathbf{x} \cdot \Delta\boldsymbol{\pi} = \sum_{a\in\mathcal{A}} x_a \sum_{A'_a \subset \mathcal{A}\setminus\{a\}} \mathbb{P}_A[A'_a] Q(\pi_*[A'_a] - \pi_a))(\mathbf{y}_a[\boldsymbol{\pi}; A'_a] - \mathbf{e}_a) \cdot \Delta\boldsymbol{\pi}.$$

with some $\mathbf{y}_a[\boldsymbol{\pi}; A'_a] \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}; A'_a])$. We have $\mathbf{y}_a[\boldsymbol{\pi}; A'_a] \cdot \Delta\boldsymbol{\pi} = \Delta\pi_*[A'_a]$; and, $\mathbf{e}_a \cdot \Delta\boldsymbol{\pi} = \Delta\pi_a$. Hence, we obtain

$$\frac{\partial G}{\partial \boldsymbol{\pi}}(\mathbf{x}, \boldsymbol{\pi})\Delta\boldsymbol{\pi} = \Delta\mathbf{x} \cdot \Delta\boldsymbol{\pi}.$$

ii) By Assumption A3, $\mathbf{g}_*(\boldsymbol{\pi})$ is constant to $\mathbf{x}$; thus, $G(\mathbf{x}, \boldsymbol{\pi}) = \mathbf{g}_*(\boldsymbol{\pi}) \cdot \mathbf{x}$ is linear in $\mathbf{x}$ with coefficient $\mathbf{g}_*(\boldsymbol{\pi})$, when $\boldsymbol{\pi}$ is fixed. Thus, by property GH-0, any $\Delta\mathbf{x} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}]$ satisfies

$$\frac{\partial G}{\partial \mathbf{x}}(\mathbf{x}, \boldsymbol{\pi})\Delta\mathbf{x} = \mathbf{g}_*(\boldsymbol{\pi}) \cdot \Delta\mathbf{x} = H(\mathbf{x}, \boldsymbol{\pi}).$$

□

A.5 Proof of Theorem 9

Proof. **G.** i) $G \geq 0$ is immediate as only the non-negative part of $\pi_*[A'] - \mathbf{x} \cdot \boldsymbol{\pi} - q$ is taken. ii) If $\mathbf{x} \in \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$, then $\pi_*[A'] - \mathbf{x} \cdot \boldsymbol{\pi} - q \leq 0$ for any $q \geq 0$ and $A' \subset \mathcal{A}$; thus, we obtain $G(\boldsymbol{\pi}, \mathbf{x}) = 0$ by Assumption Q1. On the other hand, if $\mathbf{x} \notin \Delta^{\mathcal{A}}(b_*[\boldsymbol{\pi}])$, then there is an action a such that $x_a > 0$ and $\pi_a < \pi_*$; for any optimal action $b \in b_*(\boldsymbol{\pi})$, this implies $\mathbf{x} \cdot \boldsymbol{\pi} < \pi_* = \pi_b$. Assumption A1′ guarantees that this action b is available with positive probability: there exists $A' \subset \mathcal{A}$ such that $b \in A'$ and $\mathbb{P}_A[A'] > 0$. As $\mathbf{x} \cdot \boldsymbol{\pi} < \pi_b = \pi_*[A']$ for this A', Assumption Q2 implies $\mathbb{E}_Q[\pi_*[A'] - \mathbf{x} \cdot \boldsymbol{\pi} - q] > 0$. Combining these two positives, we have $Q_*(\boldsymbol{\pi}, \mathbf{x}) \geq \mathbb{P}_A[A']\mathbb{E}_Q[\pi_*[A'] - \mathbf{x} \cdot \boldsymbol{\pi} - q] > 0$.
H. This can be proven similarly to property G.
GH-1. i) Similarly to the proof of g2 in Theorem 3, we obtain

$$\frac{\partial g_*}{\partial \boldsymbol{\pi}}(\boldsymbol{\pi}, \mathbf{x})\Delta\boldsymbol{\pi} = \sum_{A' \subset \mathcal{A}} \mathbb{P}_A[A'] Q(\pi_*[A'] - \mathbf{x} \cdot \boldsymbol{\pi})(\Delta\pi_*[A'] - \mathbf{x} \cdot \Delta\boldsymbol{\pi}).$$

Notice that the LHS coincides with $(\partial G/\partial \Delta\boldsymbol{\pi})\Delta\boldsymbol{\pi}$ and the RHS is equal to $\Delta\mathbf{x} \cdot \Delta\boldsymbol{\pi}$ for any $\Delta\mathbf{x} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}]$. Therefore, we have

$$\frac{\partial G}{\partial \boldsymbol{\pi}}(\boldsymbol{\pi}, \mathbf{x})\Delta\boldsymbol{\pi} = \frac{\partial g_*}{\partial \boldsymbol{\pi}}(\boldsymbol{\pi}, \mathbf{x})\Delta\boldsymbol{\pi} = \Delta\mathbf{x} \cdot \boldsymbol{\pi} \qquad \text{for any } \Delta\mathbf{x} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}].$$

ii) As $\mathbf{x}$ enters G only in the expectation term $\mathbb{E}_Q[\pi_*[A'] - \mathbf{x} \cdot \boldsymbol{\pi} - q]$, we have

$$\frac{\partial G}{\partial \mathbf{x}}(\boldsymbol{\pi}, \mathbf{x}) = \sum_{A' \subset \mathcal{A}} \mathbb{P}_A[A'] Q(\pi_*[A'] - \mathbf{x} \cdot \boldsymbol{\pi})(-\boldsymbol{\pi}).$$

This implies

$$\frac{\partial G}{\partial \mathbf{x}}(\boldsymbol{\pi}, \mathbf{x})\Delta\mathbf{x} = -\sum_{A' \subset \mathcal{A}} \mathbb{P}_A[A'] Q(\pi_*[A'] - \mathbf{x} \cdot \boldsymbol{\pi})\boldsymbol{\pi} \cdot \Delta\mathbf{x}.$$

As we have seen above, for any $\Delta\mathbf{x} \in \mathcal{V}(\mathbf{x})[\boldsymbol{\pi}]$, we have

$$\boldsymbol{\pi} \cdot \Delta\mathbf{x} = \sum_{A' \subset \mathcal{A}} \mathbb{P}_A[A'] Q(\pi_*[A'] - \mathbf{x} \cdot \boldsymbol{\pi})(\Delta\pi_*[A'] - \mathbf{x} \cdot \Delta\boldsymbol{\pi}).$$

Therefore, we have

$$\frac{\partial G}{\partial \mathbf{x}}(\boldsymbol{\pi}, \mathbf{x})\Delta\mathbf{x} = H(\boldsymbol{\pi}, \mathbf{x}).$$

$\square$

A.6 Proof of Theorem 8

Proof. From properties of functions G and H, we have equation (8). The quadratic term is decomposed as

$$\dot{\mathbf{x}} \cdot \frac{d\mathbf{F}}{d\mathbf{x}} \dot{\mathbf{x}} = \sum_{\substack{i \in \mathcal{A} \\ j \in \mathcal{A}}} \dot{x}_i \frac{\partial F_i}{\partial x_j}(\mathbf{x}) \dot{x}_j = \sum_{\substack{i \in S \\ j \in S}} \dot{x}_i \frac{\partial F_i}{\partial x_j}(\mathbf{x}) \dot{x}_j + \sum_{\substack{i \in S \\ j \in U}} \dot{x}_i \frac{\partial F_i}{\partial x_j}(\mathbf{x}) \dot{x}_j + \sum_{\substack{i \in U \\ j \in \mathcal{A}}} \dot{x}_i \frac{\partial F_i}{\partial x_j}(\mathbf{x}) \dot{x}_j.$$

Since $\mathbf{x}^*$ is regular ESS and $\mathbf{F}$ is C^1, the first summation in the last line is negative as long as $\sum_{a \in S} \dot{x}_a \neq 0$.

As a general algebraic property independent of specification of dynamics, Sandholm (2010a, pp.43–4) proves the existence of constant $K > 0$ such that

$$\left| \sum_{\substack{i \in U \\ j \in \mathcal{A}}} \dot{x}_i \frac{\partial F_i}{\partial x_j}(\mathbf{x}) \dot{x}_j \right|, \left| \sum_{\substack{i \in S \\ j \in U}} \dot{x}_i \frac{\partial F_i}{\partial x_j}(\mathbf{x}) \dot{x}_j \right| < K \sum_{a \in U} |\dot{x}_a|.$$

For an arbitrary action $a \in \mathcal{A}$, we can decompose the net flow $\dot{x}_a$ as[24]

$$\dot{x}_a = \sum_{b \in \mathcal{A} \setminus \{a\}} \tilde{y}_{ba} - \sum_{b \in \mathcal{A} \setminus \{a\}} \dot{x}_{ab} = \tilde{y}_{\mathcal{A}a} - \tilde{y}_{a\mathcal{A}},$$

where $\dot{x}_{ab} \in \mathbb{R}_+$ denotes the gross flow from the mass of action a-players to that of action b-players and takes a non-negative real value. Under our framework, it is exactly represented as

$$\tilde{y}_{ab} = \sum_{\substack{A'_a \subset \mathcal{A} \setminus \{a\} \\ \text{s.t. } b \in b^*(A'_a)[\boldsymbol{\pi}]}} \mathbb{P}_{Aa}(A'_a) Q(\pi_*[A'_a] - \pi_a) y_{ab}[A'_a],$$

with some $\mathbf{y}_a[A'_a] \in \Delta^A(b_*[A'_a])$ for each $A'_a \subset \mathcal{A} \setminus \{a\}$. Assumption Q1 implies that

$$\pi_i > \pi_j \quad \Longrightarrow \quad \dot{x}_{ij} = 0 \qquad \text{for all } \dot{\mathbf{x}} \in V[\boldsymbol{\pi}](\mathbf{x}). \tag{14}$$

Since $x_{ab} \geq 0$, we have

$$|\dot{x}_a| \leq \tilde{y}_{\mathcal{A}a} + \tilde{y}_{a\mathcal{A}}.$$

The aggregation over $a \in U$ yields

$$\sum_{a \in U} |\dot{x}_a| \leq \tilde{y}_{\mathcal{A}U} + \tilde{y}_{U\mathcal{A}}.$$

[24] For a set $S \subset \mathcal{A}$ and a vector $\mathbf{z} \in \mathbb{R}^A$ whose coordinates are labeled with $\mathcal{A}$, let z_S be $\sum_{a \in S} z_a$.

In a sufficiently small neighborhood of the regular ESS $\mathbf{x}^*$, say X_1^*, payoff disadvantage of actions in U compared to those in S should be maintained by continuity of $\mathbf{F}$:

$$F_s(\mathbf{x}) > F_u(\mathbf{x}) \qquad \text{for all } s \in S, u \in U, \mathbf{x} \in X^*.$$

Then, in this neighborhood, (14) implies $\dot{x}_{su} = 0$ for $s \in S, u \in U$ and thus $\tilde{y}_{SU} = 0$; this holds for any $\dot{\mathbf{x}} \in V^{\mathbf{F}}(\mathbf{x})$ with $\mathbf{x} \in X^*$. As a result, we have $\tilde{y}_{\mathcal{A}U} = \tilde{y}_{UU}$ and thus

$$\sum_{a \in U} |\dot{x}_a| \leq \tilde{y}_{UU} + \tilde{y}_{U\mathcal{A}} \leq 2\dot{x}_{U\mathcal{A}} = 2 \sum_{u \in U} \tilde{y}_{u\mathcal{A}}.$$

Here $\tilde{y}_{u\mathcal{A}}$ is the total outflow from the mass of action-u players.

From our formulation and the assumption of constant revision rate 1, it is immediate to see that

$$\tilde{y}_{a\mathcal{A}} = x_a \sum_{A'_a \subset \mathcal{A} \setminus \{a\}} \mathbb{P}_{Aa}(A'_a) Q(\pi_*[A'_a] - \pi_a) \leq x_a \qquad \text{for all } a \in \mathcal{A}, \dot{\mathbf{x}} \in V[\boldsymbol{\pi}](\mathbf{x}). \tag{15}$$

By this equation (15), the above equation reduces to

$$\sum_{a \in U} |\dot{x}_a| \leq 2 \sum_{u \in U} \tilde{y}_{u\mathcal{A}} \leq 2 \sum_{u \in U} x_u = x_U.$$

In sum, we obtain the upper bound on the quadratic term in (8):

$$\begin{aligned} \dot{\mathbf{x}} \cdot \frac{d\mathbf{F}}{d\mathbf{x}}(\mathbf{x})\dot{\mathbf{x}} &= \sum_{\substack{i \in S \\ j \in S}} \dot{x}_i \frac{\partial F_i}{\partial x_j}(\mathbf{x})\dot{x}_j + \sum_{\substack{i \in S \\ j \in U}} \dot{x}_i \frac{\partial F_i}{\partial x_j}(\mathbf{x})\dot{x}_j + \sum_{\substack{i \in U \\ j \in \mathcal{A}}} \dot{x}_i \frac{\partial F_i}{\partial x_j}(\mathbf{x})\dot{x}_j. \\ &\leq 2K \sum_{a \in U} |\dot{x}_a| \leq 4K x_U \qquad \text{for all } \mathbf{x} \in X_1^* \end{aligned} \tag{16}$$

Now we find an upper bound on $\dot{x}_U$. The above argument is again applied to obtain

$$\dot{x}_U = \tilde{y}_{UU} - \tilde{y}_{U\mathcal{A}} = -\tilde{y}_{US}.$$

By the definition of X^*, if available action set $A'_u \in \mathcal{A} \setminus U$ contains some action $s \in S$, then we have

$$F_*[A'_u](\mathbf{x}) \geq F_s(\mathbf{x}) > F_{u'}(\mathbf{x}) \qquad \text{for all } u' \in U.$$

The strict inequality implies $b_*[A'_u](\mathbf{x}) \subset S$. In a dynamic under our framework, it means that a revising agent must switch action to an action in S if the above action set A'_u. Therefore, we have

$$\tilde{y}_{uS} \geq x_u \sum_{\substack{A'_u \in \mathcal{A} \\ \text{s.t. } A'_u \cap S \neq \emptyset}} \mathbb{P}_{Au}[A'_u] Q(F_*[A'_u](\mathbf{x}) - F_u(\mathbf{x})) \cdot 1 \tag{17}$$

Since S and U are finite sets, there exists the minimum of the following minimization:

$$\hat{\pi}^* := \min_{\substack{s \in S \\ u \in U}} F_s(\mathbf{x}^*) - F_u(\mathbf{x}^*).$$

As $\mathbf{x}^*$ is a regular ESS, all of $F_s(\mathbf{x}^*) - F_u(\mathbf{x}^*)$ are (strictly) positive; the existence of the minimum then guarantees $\hat{\pi}^* > 0$. Continuity of $\mathbf{F}$ further implies that, in a sufficiently small neighborhood of $\mathbf{x}^*$, say X_2^*, we have

$$F_s(\mathbf{x}) - F_u(\mathbf{x}) \geq \min_{\substack{s \in S \\ u \in U}} F_s(\mathbf{x}) - F_u(\mathbf{x}) > \hat{\pi}^*/2.$$

By Assumption Q2, it implies

$$Q(F_s(\mathbf{x}) - F_u(\mathbf{x})) \geq Q(\hat{\pi}^*/2) > 0 \qquad \text{for all } \mathbf{x} \in X_I^*.$$

Furthermore, Assumption A1, we have

$$\bar{P}_S := \min_{u \in U} \sum_{\substack{A'_u \in \mathcal{A} \\ \text{s.t. } A'_u \cap S \neq \emptyset}} \mathbb{P}_{Au}[A'_u] > 0.$$

Plugging these two equations into (17), we have

$$\tilde{y}_{uS} \geq x_u Q(\hat{\pi}^*/2) \sum_{\substack{A'_u \in \mathcal{A} \\ \text{s.t. } A'_u \cap S \neq \emptyset}} \mathbb{P}_{Au}[A'_u] \geq x_u Q(\hat{\pi}^*/2) \bar{P}_S.$$

Summing this over all $u \in U$, we obtain

$$\dot{x}_U = -\tilde{y}_{US} \leq -Q(\hat{\pi}^*/2)\bar{P}_S x_U. \tag{18}$$

Let $\tilde{C}$ be a sufficiently large positive number such that $\tilde{C}Q(\hat{\pi}^*/2)\bar{P}_S > 4K$, say $\tilde{C} = 1 + 4K/(Q(\hat{\pi}^*/2)\bar{P}_S)$. Let $X^* == X_0^* \cap X_1^* \cap X_2^*$. With this $\tilde{C} > 0$, define function $G^* : X^* \to \mathbb{R}_+$ by

$$G^*(\mathbf{x}) := G(\mathbf{x}) + x_U.$$

As regular ESS $\mathbf{x}^*$ is an isolated Nash equilibrium in X_0^* and especially the only Nash equilibrium in X^*, we have $G(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{x}^*$ as long as $\mathbf{x} \in X^*$; $G(\mathbf{x}) > 0$ in all the other points of X^*. The second term x_U is generally nonnegative and especially zero at $\mathbf{x}^*$ (possibly at other points as well). Hence, we have

$$G^*(\mathbf{x}) \geq 0; \qquad \text{and} \qquad \forall \mathbf{x} \in X^* \; [G^*(\mathbf{x}) = 0 \;\Leftrightarrow\; \mathbf{x} = \mathbf{x}^*].$$

By combining (8), (16) and (18), we obtain

$$\dot{G}^*(\mathbf{x}) \leq H(\mathbf{x}) + (4K - \tilde{C}Q(\hat{\pi}^*/2)\bar{P}_S)x_U := H^*(\mathbf{x}).$$

Like G, we have $H \leq 0$ in general and $H(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{x}^*$ as long as $\mathbf{x} \in X^*$. By combining this fact with the above argument of x_U and the construction of $\tilde{C}$ to meet $\tilde{C}Q(\hat{\pi}^*/2)\bar{P}_S > 4K$, we have

$$H^*(\mathbf{x}) \leq 0; \qquad \text{and} \qquad \forall \mathbf{x} \in X^* \; [H^*(\mathbf{x}) = 0 \;\Leftrightarrow\; \mathbf{x} = \mathbf{x}^*].$$

These two facts on G^* and H^* mean that G^* can be used as the Lyapunov function $\mathbf{x}^*$ in the neighborhood X^*, coupled with H^* as the decaying rate function. Now Theorem 7 guarantees asymptotic stability of $\mathbf{x}^*$. □

A.7 Proof of Theorem 10

Proof. **G.** i) is immediately obtained from property G-i) of each G^p and the fact $m^p > 0$. For ii), notice that $G^{\mathcal{P}}(\mathbf{x}^{\mathcal{P}}, \boldsymbol{\pi}^{\mathcal{P}}) = 0$ is equivalent to $G^p(\mathbf{x}^p, \boldsymbol{\pi}^p) = 0$ for each $p \in \mathcal{P}$. By property G-ii) of each G^p, the latter is equivalent to $\mathbf{x}^p \in m^p \Delta^{\mathcal{A}^p}(b_*^p[\boldsymbol{\pi}])$. Therefore, $G^{\mathcal{P}}(\mathbf{x}^{\mathcal{P}}, \boldsymbol{\pi}^{\mathcal{P}}) = 0$ is equivalent to $\mathbf{x}^{\mathcal{P}} \in \times_{p \in \mathcal{P}} m^p \Delta^{\mathcal{A}^p}(b_*^p[\boldsymbol{\pi}])$.

H. Each of the two parts can be readily verified from property H, similarly to the proof of property G.

GH-0. It is immediate from aggregation:

$$H^{\mathcal{P}}(\mathbf{x}^{\mathcal{P}})[\boldsymbol{\pi}^{\mathcal{P}}] = \sum_{p \in \mathcal{P}} H^p(\mathbf{x}^p)[\boldsymbol{\pi}^p] = \sum_{p \in \mathcal{P}} \Delta \mathbf{x}^p \cdot \mathbf{g}_*^p[\boldsymbol{\pi}^p] = \Delta \mathbf{x}^{\mathcal{P}} \cdot \mathbf{g}_*^{\mathcal{P}}[\boldsymbol{\pi}^{\mathcal{P}}].$$

Property GH-0 for each population $p \in \mathcal{P}$ is used in the second equality.

GH-1. First, notice that

$$\frac{\partial G^{\mathcal{P}}}{\partial \boldsymbol{\pi}^{\mathcal{P}}}(\mathbf{x}^{\mathcal{P}}, \boldsymbol{\pi}^{\mathcal{P}})\Delta \boldsymbol{\pi}^{\mathcal{P}} = \sum_{p \in \mathcal{P}} \frac{\partial G^p}{\partial \boldsymbol{\pi}^p}(\mathbf{x}^p, \boldsymbol{\pi}^p)\Delta \boldsymbol{\pi}^p \qquad \text{for any } \Delta \boldsymbol{\pi}^{\mathcal{P}} = (\Delta \boldsymbol{\pi}^p)_{p \in \mathcal{P}} \in \mathbb{R}^{A^{\mathcal{P}}};$$

$$\frac{\partial G^{\mathcal{P}}}{\partial \mathbf{x}^{\mathcal{P}}}(\mathbf{x}^{\mathcal{P}}, \boldsymbol{\pi}^{\mathcal{P}})\Delta \mathbf{x}^{\mathcal{P}} = \sum_{p \in \mathcal{P}} \frac{\partial G^p}{\partial \mathbf{x}^p}(\mathbf{x}^p, \boldsymbol{\pi}^p)\Delta \mathbf{x}^p \qquad \text{for any } \Delta \mathbf{x}^{\mathcal{P}} = (\Delta \mathbf{x}^p)_{p \in \mathcal{P}} \in \mathbb{R}^{A^{\mathcal{P}}}.$$

This implies property GH-1 for the total: for any $\Delta \boldsymbol{\pi}^{\mathcal{P}} \in \mathbb{R}^{A^{\mathcal{P}}}$ and $\Delta \mathbf{x}^{\mathcal{P}} \in \mathcal{V}^{\mathcal{P}}(\mathbf{x}^{\mathcal{P}})[\boldsymbol{\pi}^{\mathcal{P}}]$, we have

$$\text{i)}\ \frac{\partial G^{\mathcal{P}}}{\partial \boldsymbol{\pi}^{\mathcal{P}}}(\mathbf{x}^{\mathcal{P}}, \boldsymbol{\pi}^{\mathcal{P}})\Delta \boldsymbol{\pi}^{\mathcal{P}} = \sum_{p \in \mathcal{P}} \frac{\partial G^p}{\partial \boldsymbol{\pi}^p}(\mathbf{x}^p, \boldsymbol{\pi}^p)\Delta \boldsymbol{\pi}^p = \sum_{p \in \mathcal{P}} \Delta \boldsymbol{\pi}^p \cdot \Delta \mathbf{x}^p = \Delta \boldsymbol{\pi}^{\mathcal{P}} \cdot \Delta \mathbf{x}^{\mathcal{P}},$$

$$\text{ii)}\ \frac{\partial G^{\mathcal{P}}}{\partial \mathbf{x}^{\mathcal{P}}}(\mathbf{x}^{\mathcal{P}}, \boldsymbol{\pi}^{\mathcal{P}})\Delta \mathbf{x}^{\mathcal{P}} = \sum_{p \in \mathcal{P}} \frac{\partial G^p}{\partial \mathbf{x}^p}(\mathbf{x}^p, \boldsymbol{\pi}^p)\Delta \mathbf{x}^p = \sum_{p \in \mathcal{P}} H^p(\mathbf{x}^p, \boldsymbol{\pi}^p) = H^{\mathcal{P}}(\mathbf{x}^{\mathcal{P}}, \boldsymbol{\pi}^{\mathcal{P}})$$

Property GH-1 for each population $p \in \mathcal{P}$ is used in the second equality on each of the two lines. □

A.8 Proof of Theorem 11

Proof. Fix $\mathbf{x}^{\mathcal{P}}, \mathbf{y}^{\mathcal{P}} \in \mathcal{X}^{\mathcal{P}} = (\Delta^A)^P$ arbitrarily. Let $\bar{\mathbf{x}}, \bar{\mathbf{y}}$ be their aggregates: $\bar{\mathbf{x}} = \sum_{q\in\mathcal{P}} \mathbf{x}^q$ and $\bar{\mathbf{y}} = \sum_{q\in\mathcal{P}} \mathbf{y}^q$.

Notice that, under additive separability, we have

$$\mathbf{F}^p(\mathbf{y}^{\mathcal{P}}) - \mathbf{F}^p(\mathbf{x}^{\mathcal{P}}) = \mathbf{F}^0(\bar{\mathbf{y}}) - \mathbf{F}^0(\bar{\mathbf{x}}).$$

Therefore, we have

$$(\mathbf{y}^{\mathcal{P}} - \mathbf{x}^{\mathcal{P}}) \cdot (\mathbf{F}^{\mathcal{P}}(\mathbf{y}^{\mathcal{P}}) - \mathbf{F}^{\mathcal{P}}(\mathbf{x}^{\mathcal{P}})) = \sum_{p\in\mathcal{P}} (\mathbf{y}^p - \mathbf{x}^p) \cdot (\mathbf{F}^p(\mathbf{y}^{\mathcal{P}}) - \mathbf{F}^p(\mathbf{x}^{\mathcal{P}}))$$
$$= \left\{ \sum_{p\in\mathcal{P}} (\mathbf{y}^p - \mathbf{x}^p) \right\} \cdot (\mathbf{F}^0(\bar{\mathbf{y}}) - \mathbf{F}^0(\bar{\mathbf{x}})) = (\bar{\mathbf{y}} - \bar{\mathbf{x}}) \cdot (\mathbf{F}^0(\bar{\mathbf{y}}) - \mathbf{F}^0(\bar{\mathbf{x}})).$$

Hence, $\mathbf{F}^{\mathcal{P}}$ is contractive (resp., strictly contractive) if so is $\mathbf{F}^0$. □

A.9 Proof of Theorem 12

Proof. (11) implies the negative definiteness of $d\mathbf{F}^{\mathcal{P}}/d\mathbf{x}^{\mathcal{P}}$ with respect to $T\mathcal{X}^{\mathcal{P}}$. Thanks to Assumption F2-ii) and continuity of $d\mathbf{F}^{\mathcal{P}}_\delta/d\mathbf{x}^{\mathcal{P}}$ in δ, the negative definiteness remains to hold for $d\mathbf{F}^{\mathcal{P}}_\delta/d\mathbf{x}^{\mathcal{P}}$ as long as δ is sufficiently small. Therefore, for such δ, $\mathbf{F}^{\mathcal{P}}$ is strictly contractive. Then, i) there uniquely exists a Nash equilibrium and ii) the Nash equilibrium is globally asymptotically stable under any dynamics in our framework.

With the unique existence of Nash equilibrium at each parameter value, continuity of $\mathbf{F}^{\mathcal{P}}_\delta$ in parameter δ also implies that a Nash equilibrium is represented as a continuous function of δ, say $\mathbf{x}^*(\delta)$. Notice that $\bar{F}^{\mathcal{P}}_\delta(\mathbf{x};\delta)$ is continuous in $\mathbf{x}$ and δ. Therefore, $\Delta\bar{F}^{\mathcal{P}}(\delta) := \bar{F}^{\mathcal{P}}_\delta(\mathbf{x}^*(\delta)) - \max_{\mathbf{x}^{\mathcal{P}}\in\mathcal{X}^{\mathcal{P}}} \bar{F}^{\mathcal{P}}_\delta(\mathbf{x}^{\mathcal{P}})$ is continuous in δ. As $\bar{F}^{\mathcal{P}}_0 \equiv \bar{F}^{\mathcal{P}}$ is maximized at $\mathbf{x}^*(0)$, we have $\Delta\bar{F}^{\mathcal{P}}(0) = 0$. Therefore, continuity of $\Delta\bar{F}^{\mathcal{P}}$ implies that

$$\forall\varepsilon > 0\ \exists\bar{\delta}(\varepsilon) > 0 \quad |\delta - 0| < \delta(\varepsilon) \quad \Longrightarrow \quad |\Delta\bar{F}^{\mathcal{P}}(\delta) - 0| < \varepsilon.$$

With global asymptotic stability of $\mathbf{x}^*(\delta)$, we verify the theorem. □

References

BENAÏM, M., AND O. RAIMOND (2010): "A class of self-interacting processes with applications to games and reinforced random walks," *SIAM Journal on Control and Optimization*, 48(7), 4707–4730.

BOMZE, I. M., AND J. W. WEIBULL (1995): "Does Neutral Stability Imply Lyapunov Stability?," *Games and Economic Behavior*, 11(2), 173–192.

BRAVO, M., AND M. FAURE (2015): "Reinforcement learning with restrictions on the action set," *SIAM Journal on Control and Optimization*, 53(1), 287–312.

CRESSMAN, R. (1997): "Local stability of smooth selection dynamics for normal form games," *Mathematical Social Sciences*, 34(1), 1–19.

FOX, M. J., AND J. S. SHAMMA (2013): "Population games, stable games, and passivity," *Games*, 4(4), 561–583.

FUJISHIMA, S. (2012): "Evolutionary Implementation with Estimation of Unknown Externalities," mimeo, Washington University in St. Louis.

FUJIWARA-GREVE, T., AND M. OKUNO-FUJIWARA (2009): "Voluntarily Separable Repeated Prisoner's Dilemma," *Review of Economic Studies*, 76(3), 993–1021.

——— (2016): "Diverse Behavior Patterns in a Symmetric Society with Voluntary Partnerships," mimeo, Keio University and Musashino University.

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

HOFBAUER, J., P. SCHUSTER, AND K. SIGMUND (1979): "A note on evolutionary stable strategies and game dynamics," *Journal of Theoretical Biology*, 81(3), 609–612.

LAHKAR, R., AND W. H. SANDHOLM (2008): "The projection dynamic and the geometry of population games," *Games and Economic Behavior*, 64(2), 565–590.

NAGURNEY, A., AND D. ZHANG (1997): "Projected dynamical systems in the formulation, stability analysis, and computation of fixed-demand traffic network equilibria," *Transportation Science*, 31(2), 147–158.

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.

——— (2013): "Population Games and Deterministic Evolutionary Dynamics," in *Handbook of Game Theory and Economic Applications*, ed. by H. P. Young, and S. Zamir, vol. 4. Elsevier., forthcoming.

SANDHOLM, W. H., E. DOKUMACI, AND R. LAHKAR (2008): "The projection dynamic and the replicator dynamic," *Games and Economic Behavior*, 64(2), 666–683.

SAWA, R., AND D. ZUSAI (2014): "Evolutionary imitative dynamics with population-varying aspiration levels," *Journal of Economic Theory*, 154, 562–577.

——— (2016): "Best response dynamics in multitask environments," mimeo, University of Tsukuba and Temple University, `http://sites.temple.edu/zusai/research/multiBRD/`.

TAYLOR, P., AND L. JONKER (1978): "Evolutionarily Stable Strategies and Game Dynamics," *Mathematical Biosciences*, 40, 145–156.

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. (2015): "Tempered best response dynamics," mimeo, Temple University, `http://sites.temple.edu/zusai/research/tbrd/`.