# Stable Mixing in Hawk–Dove Games under Best Experienced Payoff Dynamics*

Srinivas Arigapudi†, Yuval Heller‡

August 30, 2024

**Abstract**

The hawk–dove game admits two types of equilibria: an asymmetric pure equilibrium, in which players in one population play "hawk" and players in the other population play "dove," and an inefficient symmetric mixed equilibrium, in which hawks are frequently matched against each other. The existing literature shows that populations will converge to playing one of the pure equilibria from almost any initial state. By contrast, we show that plausible dynamics, in which agents occasionally revise their actions based on the payoffs obtained in a few past interactions, can give rise to the opposite result: convergence to one of the interior stationary states.

**Keywords:** hawk–dove game, chicken game, learning, evolutionary stability, best experienced payoff dynamics. **JEL codes:** C72, C73.

# 1. Introduction

The hawk–dove game is widely used to study situations of conflict between strategic participants.[1] As a simple motivating example, consider a situation in which a buyer (Player 1) and a seller (Player 2) bargain over the price of an asset (e.g., a house). Each

*This paper contains some parts of an unpublished working paper titled "Sampling Dynamics and Stable Mixing in Hawk–Dove Games." The authors thank Augusto Santos and Amnon Schreiber for various helpful comments and suggestions.

†Department of Economic Sciences, Indian Institute of Technology Kanpur, India. e-mail: arigapudi@iitk.ac.in; website: https://sites.google.com/view/srinivasarigapudi/home.

‡Department of Economics, Bar-Ilan University, Ramat Gan 5290002, Israel. e-mail: yuval.heller@biu.ac.il; website: https://sites.google.com/site/yuval26/.

[1] A few examples of applications of the hawk–dove game (also known as the chicken game; see, e.g., Rapoport and Chammah, 1966; Aumann, 1987) are: provision of public goods (Lipnowski and Maital, 1983), nuclear deterrence between superpowers (Brams and Kilgour, 1987; Dixit et al., 2019), industrial disputes (Bornstein et al., 1997), bargaining problems (Brams and Kilgour, 2001), conflicts between countries over contested territories (Baliga and Sjöström, 2012, 2020), and task allocation among members of a team (Herold and Kuzmics, 2020).

Table 1: Payoff Matrix of the Standard Hawk–Dove Game ($g \in (0,1)$)

| | | Player 2: $h$ | Player 2: $d$ |
|---|---|---|---|
| Player 1 | $h$ | $0, 0$ | $1+g, 1-g$ |
| | $d$ | $1-g, 1+g$ | $1, 1$ |

player has two possible bargaining strategies (actions): insisting on a more favorable price (referred to as being a "hawk"), or agreeing to a less favorable price in order to close the deal (being a "dove"). The payoffs of the game are presented in Table 1. Two doves agree on a price that is equally favorable to both sides, and obtain a relatively high payoff, which is normalized to 1. A hawk obtains a favorable price when being matched with a dove, which yields her an additional gain of $g \in (0,1)$, at the expense of her dovish opponent.[2] Finally, two hawks obtain the lowest payoff of 0, due to a substantial probability of bargaining failure.[3] Observe that large values of $g$ (close to 1) correspond to environments that are advantageous to hawks (i.e., being a hawk yields a higher expected payoff against an opponent who might play either action with equal probability), small values of $g$ correspond to environments that are advantageous to doves, and values of $g$ that are close to 0.5 correspond to approximately balanced environments.

The hawk–dove game admits three Nash equilibria: two asymmetric pure equilibria, and an inefficient symmetric mixed (interior) equilibrium. In the pure equilibria (in which one of the players plays hawk while the opponent plays dove), all conflicts are avoided at the cost of inequality, as the payoff of the hawkish player is substantially higher than that of the dovish opponent. By contrast, in the symmetric interior equilibrium both players obtain the same expected payoff, yet this payoff is relatively low due to the positive probability of a conflict arising between two hawks.

A natural question is to ask which equilibrium is more likely to obtain. Standard game theory is not helpful in answering this question, as all these Nash equilibria satisfy all the standard refinements (e.g., perfection). By contrast, the dynamic (evolutionary) approach can yield sharp predictions (for textbook expositions, see Weibull, 1997; Sandholm, 2010).

Consider a setup in which pairs of agents from two infinite populations are repeatedly matched at random times (each such match of an agent from population 1 is against an

[2] Our formal model studies a broader class of generalized hawk–dove games, in which the gain of a hawkish player might differ from the loss of the dovish opponent (as shown in Table 2 in Section 2).

[3] Our one-parameter payoff matrix is equivalent to the commonly used two-parameter payoff matrix (Maynard Smith, 1982), in which a dove obtains $\frac{V}{2}$ against another dove and 0 against a hawk, and a hawk obtains $\frac{V-C}{2}$ against another hawk and $V$ against a dove. Specifically, our one-parameter matrix is obtained from the two-parameter matrix by the affine transformation of adding the constant $\frac{C-V}{2}$ and dividing all payoffs by $\frac{C}{2}$, followed by substituting $g \equiv \frac{V}{C}$.

opponent from population 2).[4] Agents occasionally die (or, alternatively, agents occasionally receive opportunities to revise their actions). New agents observe some information about the aggregate behavior and the payoffs, and use this information to choose the action they will play in all future encounters. We are interested in characterizing the stable rest points of such revision dynamics, which can be used as an equilibrium refinement.

Most existing models assume that the revision dynamics are *monotone* (also known as sign-preserving) with respect to the payoffs: the frequency of the strategy that yields the higher payoff (of the two feasible strategies) increases. A key result in evolutionary game theory is that in a hawk–dove game, all monotone (two-population) revision dynamics converge to the asymmetric pure equilibria from almost any initial state (henceforth, *global convergence*; see Maynard Smith and Parker, 1976, for the classic analysis, Maynard Smith, 1982, for the textbook presentation, Sugden, 1989, for the economic implications, and Oprea et al., 2011, for the general dynamic result). Thus, the existing literature predicts that an efficient convention will emerge in which trade always occurs and most of the surplus goes to one side of the market. Casual observation suggests that this prediction might not fit well the behavior in situations such as the motivating example, in which the surplus of trade is typically divided relatively equally between the two sides of the market, and in which bargaining frequently fails.

In many applications, precise information about the aggregate behavior in the population may be difficult or costly to obtain. In such situations, new agents have to infer the aggregate behavior in the population from a small sample of other players. Plausible revision dynamics in such situations are the *best experienced payoff dynamics* (BEP dynamics for short; Sandholm et al., 2020). Under these dynamics, a new agent in the population tries each of the feasible actions $k$ times, using it each time against a newly drawn set of opponents from the opponent population. The agent then chooses the action that yielded the highest mean payoff in these trials, employing some tie-breaking rule if more than one action yields the highest mean payoff.

We analyze the BEP dynamics in the hawk–dove game. In our analysis, we allow agents to have heterogeneous sample sizes (i.e., each new agent is endowed with a sample size of $k$ that is randomly chosen from an exogenous distribution). It is relatively straightforward to show that these dynamics admit at least three stationary states: two asymmetric pure states, and an inefficient symmetric interior state.[5]

[4]Our paper focuses on two-population dynamics in which players condition their play on their role in the game. By contrast, in a one-population model, players cannot condition their play on their role in the game. The predictions of the one-population model are discussed in Remark 1.

[5]Although the symmetric stationary state does not coincide with the symmetric Nash equilibrium, they share similar qualitative properties: namely, symmetry between the two populations, and inefficiency induced by frequent matching of two hawks.

We show that the BEP dynamics can yield qualitatively different results compared to monotone dynamics in the hawk–dove game. Specifically, our first main result (Theorem 1) presents a simple condition for BEP dynamics to induce the opposite result (relative to monotone dynamics), namely, convergence to one of the inefficient interior stationary states from any initial interior state. This simple condition holds for various distributions of small sample sizes. In particular, if $g > \frac{1}{3}$ it holds whenever at least half of the players have sample size 2 (regardless of the sample size of the remaining players in the population).

Our second main result (Theorem 2) shows that under the BEP dynamics, if $g$ is not too far from one, then many populations with relatively small sample sizes converge to one of the inefficient interior stationary states from any sufficiently close initial state. Specifically, the condition of Theorem 2 holds (1) in any homogeneous population in which the uniform sample size of all agents is at most 20, and (2) in any heterogeneous population in which the uniform sample size of all agents is at most five.

Taken together, our results show that when some agents have limited information about the aggregate behavior of the population (which seems plausible in various real-life applications, such as the motivating example of buyers and sellers of houses), then an egalitarian, yet inefficient, convention may arise in which bargaining frequently fails.

Our model can help to explain some stylized empirical facts suggesting inefficiency in bargaining in real-life markets. Koster and Rouwendal (2023) empirically analyzed the Dutch housing market. They found that many sellers tend to set too low list prices, which can be consistent with profit maximization only if the sellers have high annual discount factors of up to 50. This behavior is consistent with an interior stable state in our model in which some sellers apply the BEP dynamics with small samples. Larsen (2021) empirically studied the efficiency of bargaining following wholesale used- car auctions in which the highest bid is lower than the reserve price. Larsen found that bargaining fails in about $35\%$ of these cases, which is consistent with an inefficient interior stable stationary state in our model. The data suggests that this inefficiency results in substantial losses ($12\% - 23\%$ of ex-post gains from trade). Moreover, Larsen's analysis suggests that only a small part of this loss is due to incomplete information constraints.

An experimental setup that is likely to induce behavior close to the predictions of the BEP dynamics is the "black box" setting, in which players do not know the game's structure and observe only their realized payoffs (see, e.g., Nax and Perc, 2015; Nax et al., 2016; Burton-Chellew et al., 2017). Arguably, the black box setting is relevant to many real-life interactions. To the best of our knowledge, hawk–dove games have not been experimentally tested in a black box setting. Our model predicts that if subjects play (two-population) hawk–dove games with random matching and a sufficiently high parameter $g$ in the black box setting, then the behavior would converge to an inefficient symmetric interior state.

The paper proceeds as follows: Section 1.1 presents the related literature and the main contribution of this paper. Section 2 introduces our model. Section 3 presents the results and Section 4 concludes.

### 1.1 Literature and Contribution

BEP dynamics were pioneered by Osborne and Rubinstein (1998) and Sethi (2000) and later generalized in various respects by Sandholm et al. (2020). They have been used in a variety of applications, including price competition with boundedly rational consumers (Spiegler, 2006), common-pool resources (Cárdenas et al., 2015), contributions to public goods (Mantilla et al., 2018), centipede games (Sandholm et al., 2019), finitely repeated games (Sethi, 2021), the prisoner's dilemma (Arigapudi et al., 2021), the traveler's dilemma (Berkemer et al., 2023), and the trust game (Arigapudi and Lahkar, 2024). Izquierdo and Izquierdo (2022) provide simple algorithms to determine the stability or instability of strict Nash equilibria under the BEP dynamics (see also Izquierdo and Izquierdo, 2023).

A related modeling approach studied in the literature are the sampling best response dynamics (see, e.g., Sandholm, 2001; Oyama et al., 2015; Osborne and Rubinstein, 2003; Salant and Cherry, 2020; Sawa and Wu, 2023; Danenberg and Spiegler, 2022.) Under these dynamics, the assumption is that revising agents in the population use information from samples of opponents' play to form point estimates of the population distribution of actions, and then choose an action that is a best response to this estimate. Under the sampling best response dynamics, it is assumed that the revising agents know the payoffs of the underlying game that is being played in the population. By contrast, under BEP dynamics, the restrictions on revising agents' informational and computational skills are significantly less demanding. Revising agents under the BEP dynamics do not know the payoffs of the game they are playing, or even that they are playing a game; only payoff experiences count.

**Main Contribution:** It is well known that interior stationary states in multiple-population games cannot be asymptotically stable[6] under the commonly studied replicator dynamics (see, e.g., Sandholm, 2010, Theorem 9.1.6). Recently, Arigapudi et al. (2023) showed

[6] A stationary state is said to be asymptotically stable if it attracts all trajectories starting from nearby states. A stationary state is globally stable if the population converges to it starting from any initial interior state. See Definitions 5 and 6.

that in a two-population coordination game,[7] interior stationary states can be asymptotically stable under the sampling best response dynamics only if there is heterogeneity in the sample sizes: some agents have accurate information based on large samples of the opponents' aggregate behavior, while other agents rely on anecdotal evidence induced by small samples. A key contribution of this paper is to show that in a two-population hawk–dove game, interior stationary states can be asymptotically stable under BEP dynamics even when agents in the population have homogeneous sample sizes.

# 2. Model

## 2.1 The Hawk–Dove Game

Let $G = \{A, u\}$ denote a symmetric two-player hawk–dove game, where:

1. $A = \{h, d\}$ is the set of actions of each player, and
2. $u : A^2 \rightarrow \mathbb{R}$ is the payoff function of each player.

Let $i \in \{1, 2\}$ be an index referring to one of the players, and let $j = \{1, 2\} \setminus \{i\}$ be an index referring to the opponent. We interpret action $h$ as the hawkish (more aggressive) action and $d$ as the dovish action. The payoff matrix $u(\cdot, \cdot)$ of a generalized hawk–dove game is given in Table 2. When both agents are dovish, they obtain a relatively high payoff, which is normalized to 1. When both agents are hawkish, they obtain their lowest feasible payoff, which is normalized to 0. Finally, when one of the players is hawkish and her opponent is dovish, the hawkish player *gains* $g \in (0, 1)$ (relative to the payoff of 1 obtained by two dovish players), while her dovish opponent *loses*[8] $l \in (0, 1)$. The game admits three Nash equilibria: two asymmetric pure Nash equilibria: $(h, d)$ and $(d, h)$, and a symmetric interior Nash equilibrium in which each player plays $h$ with probability $\frac{g}{1+g-l}$, and obtains a relatively low expected payoff of $\frac{(1+g)(1-l)}{1+g-l} < 1$.

An important special subclass is the *standard* hawk–dove games, in which $g = l$ (see Table 1); i.e., the gain of the hawkish player is equal to the loss of her dovish opponent.

[7] Notice that a two-population hawk–dove game can be modeled as a two-population coordination game after relabeling of actions.

[8] Herold and Kuzmics (2020) extend the domain to one in which the assumption of $g, l \in (0, 1)$ is replaced with the weaker assumption of $g > 0$, $l < 1$, and $l + g > 0$. All of our results hold in this extended domain.

Table 2: Payoff Matrix of a Generalized Hawk–Dove Game $g, l \in (0,1)$

| | | Player 2 $h$ | Player 2 $d$ |
|---|---|---|---|
| Player 1 | $h$ | $0, 0$ | $1+g, 1-l$ |
| | $d$ | $1-l, 1+g$ | $1, 1$ |

## 2.2 Evolutionary Process

We assume that there are two unit-mass continuums of agents (e.g., buyers and sellers) and that agents in population 1 are randomly matched with agents in population 2. Aggregate behavior in the populations at time $t \in \mathbb{R}^+$ is described by a *state* $\mathbf{p}(t) = (p_1(t), p_2(t)) \in [0,1]^2$, where $p_i(t)$ represents the share of agents playing the hawkish action $h$ at time $t$ in population $i$. We extend the payoff function $u$ to states (which have the same representation as mixed strategy profiles) in the standard linear way. Specifically, $u(p_i, p_j)$ denotes the average payoff of population $i$ (in which a share $p_i$ of the population plays $h$) when randomly matched against population $j$ (in which a share $p_j$ plays $h$). With a slight abuse of notation, we use $d$ (resp., $h$) to denote a degenerate population in which all of its agents play action $d$ (resp., $h$). A state $\mathbf{p} = (p_1, p_2)$ is *symmetric* if $p_1 = p_2$. A state $\mathbf{p} = (p_1, p_2)$ is *interior* (or *mixed*) if $p_1, p_2 \in (0,1)$.

Agents occasionally die and are replaced by new agents (equivalently, agents occasionally receive opportunities to revise their actions). Let $\delta > 0$ denote the death rate of agents in each population, which we assume to be independent of the currently used actions. It turns out that $\delta$ does not have any effect on the dynamics, except to multiply the speed of convergence by a constant. The evolutionary process is represented by a continuous function $\mathbf{w} : [0,1]^2 \to [0,1]^2$, which describes the frequency of new agents in each population who adopt action $h$ as a function of the current state. That is, $w_i(\mathbf{p})$ describes the share of new agents of population $i$ who adopt action $h$, given state $\mathbf{p}$. Thus, the instantaneous change in the share of agents of population $i$ that play hawk is given by $\dot{p}_i = \delta \cdot (w_i(\mathbf{p}) - p_i)$.

*Remark 1.* Our *two-population* dynamics fit situations in which each player can condition her play on her role in the game (being Player 1 or Player 2). Common examples of such situations are (1) when sellers are matched with buyers, as in the motivating example, and (2) when each player observes if she has arrived slightly earlier or slightly later at a contested resource (Maynard Smith, 1982). The two-population dynamics essentially induce the same results as one-population dynamics over a larger game with $2 \times 2$ "role-conditioned" actions (see, e.g., Weibull, 1997, end of Section 5.1.1): being a hawk in both roles, being a dove in both roles, being a hawk as Player 1 and a dove as Player 2 (the "bourgeois" strategy of Maynard Smith, 1982), and being a hawk as Player 2 and a dove as Player 1.

By contrast, in *one-population* dynamics of the original two-action game, an agent cannot condition her play on her role. It is well known that all monotone one-population dynamics converge to the unique interior Nash equilibrium in hawk–dove games (see, e.g., Weibull, 1997, Section 4.3.2). It is relatively straightforward to establish that one-population BEP dynamics lead to qualitatively similar results (convergence is to a somewhat different interior state than in the interior Nash equilibrium, but the comparative statics with respect to the payoff parameters remain similar).

### 2.2.1 Monotone Dynamics

The most widely studied dynamics are those that are monotone with respect to the payoffs. A dynamic is monotone if the share of agents playing an action increases iff the action yields a higher payoff than the alternative action.[9,10]

**Definition 1.** *The dynamic* $\boldsymbol{w} : [0,1]^2 \to [0,1]^2$ *is* monotone *if for any player* $i$*, any interior* $p_i \in (0,1)$*, and any* $p_j \in [0,1]$ *, it holds that* $\dot{p}_i > 0 \;\Leftrightarrow\; u(h,p_j) > u(d,p_j)$ *.*

Oprea et al. (2011) show that under monotone dynamics, from almost any initial state, the populations converge to one of the two asymmetric pure equilibria in which one population always plays $h$ and the other population always plays $d$ (generalizing the seminal analysis of Maynard Smith and Parker, 1976).

### 2.2.2 Best Experienced Payoff Dynamics

In what follows, we study a plausible nonmonotone dynamic, in which new agents base their choice on inference from small samples.

**Distribution of sample sizes:** We allow heterogeneity in the sample sizes used by new agents. Let $\theta \in \Delta(\mathbb{Z}_+)$ denote the distribution of sample sizes of new agents. We assume that $\theta$ has a finite support. A share of $\theta(k)$ of the new agents have a sample of size $k$. Let supp $(\theta)$ denote the support of $\theta$, and let $\max(\text{supp}(\theta))$ denote the maximal sample size in the support of $\theta$. If there exists some $k$ for which $\theta(k) = 1$, then we use $k$ to denote the degenerate (homogeneous) distribution $\theta \equiv k$.

[9]In games with more than two actions, there are various definitions that capture different aspects of monotonicity. All these definitions coincide for two-action games. In particular, Definition 1 coincides in two-action games with Weibull's (1997, Section 5.5) textbook definitions of payoff monotonicity, payoff positivity, sign preserving, and weak payoff positivity.

[10]The best-known example of payoff monotone dynamics is the standard replicator dynamic (Taylor, 1979), which is given by $\dot{p}_i = w_i(\mathbf{p}) - p_i = p_i\left(u(h,p_j) - u(p_i,p_j)\right)$.

**Definition 2.** *An* environment *is a tuple $E = (g, l, \theta)$, where $g, l \in (0, 1)$ describes the underlying hawk–dove game, and $\theta \in \Delta(\mathbb{Z}_+)$ describes the distribution of sample sizes.*

A new $k$-agent observes for each of her feasible actions the mean payoff obtained by playing this action in $k$ interactions (with each play of each action being against a newly drawn opponent), and then chooses the action whose mean payoff was highest. One possible interpretation of these observations is that each new $k$-agent tests each of the available actions $k$ times, and then adopts for the rest of her life the action with the highest mean payoff during the testing phase. We assume that when a tie occurs, the new agent adopts action $d$ (the results are qualitatively the same for any tie-breaking rule).

We refer to the sample against which action $h$ (resp., $d$) is tested as the $h$-sample (resp., $d$-sample). Let $X(k, p_j), Y(k, p_j) \sim Bin(k, p_j)$ denote two iid random variables with a binomial distribution with parameters $k$ and $p_j$. The random variable $X(k, p_j)$ (resp., $Y(k, p_j)$) is interpreted as the number of times in which the opponents have played action $h$ in the $h$-sample (resp., $d$-sample). Observe that the sum of payoffs of playing action $h$ (resp., $d$) against its $h$-sample (resp., $d$-sample) is $(1+g)(k - X(k, p_j))$ (resp., $k - lY(k, p_j)$). This implies that action $h$ has the highest mean payoff iff $(1+g)(k - X(k, p_j)) > k - lY(k, p_j) \Leftrightarrow (1+g)X(k, p_j) < gk + lY(k, p_j)$. Thus, the BEP dynamics in the environment $(g, l, \theta)$ are given by $\dot{p}_i = w_\theta(p_j) - p_i$ for $i, j = \{1, 2\}$ and $j \neq i$, where

$$w_\theta(p_j) = \sum_{k \in \text{supp}(\theta)} \theta(k) \cdot \mathbb{P}((1+g)X(k, p_j) < gk + lY(k, p_j)). \tag{2.1}$$

## 2.3 Standard Definitions of Dynamic Stability

We present the standard definitions of dynamic stability that we use in the paper (see, e.g., Weibull, 1997, Chapter 5).

A state is said to be stationary if it is a rest point of the dynamics.

**Definition 3.** *State $\mathbf{p}^* \in [0, 1]^2$ is a* stationary state *if $w_i(\mathbf{p}^*) = p_i^*$ for each $i \in \{1, 2\}$.*

Let $\mathcal{E}(w)$ denote the set of stationary states of $w$, i.e., $\mathcal{E}(w) = \{\mathbf{p}^* \mid w_i(\mathbf{p}^*) = p_i^*\}$. Under monotone dynamics, an interior (mixed) state $\mathbf{p}^* \in (0, 1)^2$ is a stationary state iff it is a Nash equilibrium (Weibull, 1997, Prop. 4.7). By contrast, under nonmonotone dynamics, the two notions differ.

A state is Lyapunov stable if a population beginning near it remains close, and it is asymptotically stable if, in addition, it eventually converges to it. A state is (almost) globally stable if the population converges to it from any initial interior state. A state is unstable if it is not Lyapunov stable. It is well known (see, e.g., Weibull, 1997, Section 6.4) that every Lyapunov stable state must be a stationary state. Formally:

**Definition 4.** A stationary state $\mathbf{p}^* \in [0,1]^2$ is *Lyapunov stable* if for every neighborhood $U$ of $\mathbf{p}^*$ there is a neighborhood $V \subseteq U$ of $\mathbf{p}^*$ such that if the initial state $p(0) \in V$, then $\mathbf{p}(t) \in U$ for all $t > 0$. A state is *unstable* if it is not Lyapunov stable.

**Definition 5.** A stationary state $\mathbf{p}^* \in [0,1]^2$ is *asymptotically stable* if it is Lyapunov stable and there is some neighborhood $U$ of $\mathbf{p}^*$ such that all trajectories initially in $U$ converge to $\mathbf{p}^*$, i.e., $\mathbf{p}(0) \in U$ implies $\lim_{t\to\infty} \mathbf{p}(t) = \mathbf{p}^*$.

**Definition 6.** A stationary state $\mathbf{p}^* \in [0,1]^2$ is *globally stable* if it is Lyapunov stable and, for any neighborhood $U$ of $\mathbf{p}^*$, all trajectories initially in $U$ converge to $\mathbf{p}^*$.

# 3. Results

## 3.1 Global Convergence to Inefficient Interior States

As discussed in the Introduction, various papers in the literature show that in hawk–dove games, under many evolutionary dynamics eventually all agents in one population play hawk while all agents in the other population play dove. In this section, we fully characterize the conditions for which the opposite result holds under the BEP dynamics; i.e., the populations converge from almost any initial state to inefficient interior states in which players from both populations play $h$ with positive probability.

The following lemma characterizes when a single appearance of a rare action in a new agent's sample can change the agent's behavior.

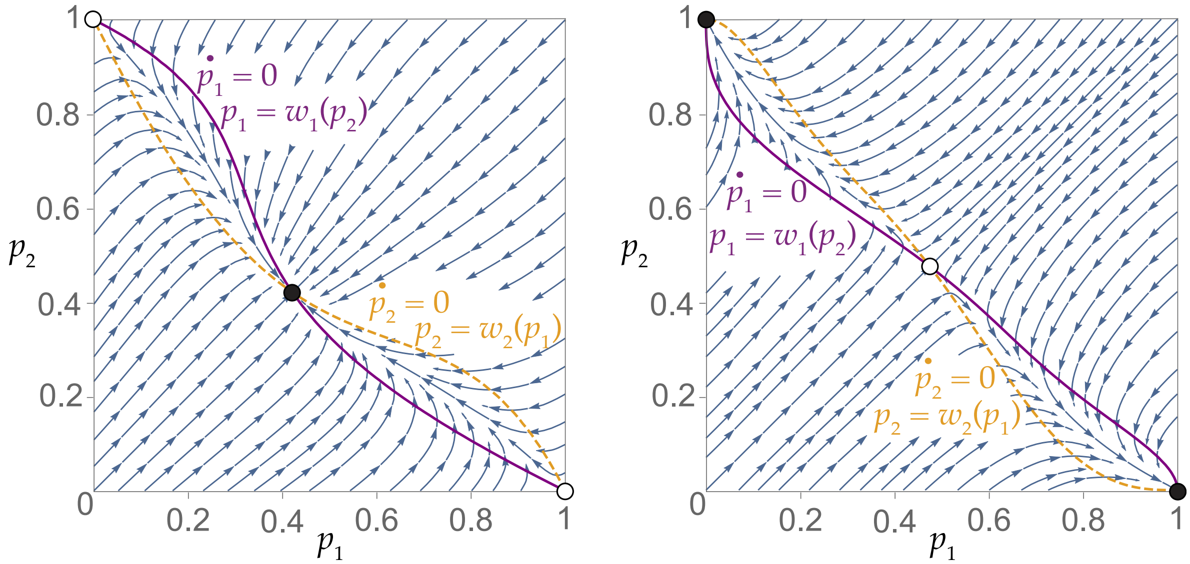
**Lemma 1.** *Consider a new agent in population $i$ with a sample size of $k$. Under the BEP dynamics, (I) an $h$-sample with a single $d$ induces a strictly higher mean payoff than a $d$-sample with no $d$-s iff $k < \frac{1+g}{1-l}$ and (II) a $d$-sample with no $h$-s induces a weakly higher mean payoff than an $h$-sample with a single $h$ iff $k \le \frac{1+g}{g}$.*

*Proof.* (I) The sum of payoffs of an $h$-sample with a single $d$ is equal to $1+g$. The sum of payoffs of a $d$-sample with no $d$ is equal to $k \cdot (1-l)$. The former sum is higher than the latter iff $k < \frac{1+g}{1-l}$. (II) The sum of payoffs of a $d$-sample with no $h$-s is equal to $k$. The sum of payoffs of an $h$-sample with a single $h$ is equal to $(k-1)(1+g)$. The former sum is higher than the latter iff $k < \frac{1+g}{g}$. □

Lemma 1 allows us to define the upper bounds on the sample size in which a single appearance of a rare action can change the behavior of a new agent.

**Definition 7.** *Let $m_h = \frac{1+g}{1-l}$, $m_d = \frac{1+g}{g}$.*

Figure 1: Phase Plots Illustrating Theorem 1




The figure illustrates the phase plots of BEP dynamics for two environments: left panel, $g = l = 0.4, k = 2$; right panel, $g = l = 0.4, k = 5$. The solid purple (resp., dashed orange ) curve of $w_1(p_2)$ (resp., $w_2(p_1)$) shows the states for which $\dot{p}_1 = 0$ (resp., $\dot{p}_2 = 0$ ). The intersection points of these curves are the stationary states. A solid (resp., hollow) dot represents an asymptotically stable (resp., unstable) stationary state. The environment in the left panel satisfies the condition of Case (1) of Theorem 1 ($\mathbb{E}_{<m_h}(\theta) \cdot \mathbb{E}_{\leq m_d}(\theta) = 2 \cdot 2 = 4 > 1$), which implies the instability of the pure stationary states. The environment in the right panel satisfies the condition of Case (2) of Theorem 1 ($\mathbb{E}_{<m_h}(\theta) \cdot \mathbb{E}_{\leq m_d}(\theta) = 0 \cdot 0 < 1$), which implies the asymptotic stability of the pure states.

The parameter $m_h$ is the upper bound on the sample size for which a single appearance of $d$ in the sample, when all other sampled actions are $h$, induces a new agent to adopt action $h$. Similarly, $m_d$ is the upper bound on the sample size for which a single appearance of $h$ in the sample, when all other sampled actions are $d$, can induce a new agent to adopt action $d$.

We conclude this subsection by presenting a definition of *$m$-bounded expectation* of a probability distribution with support on the set of positive integers. It is the expected value of the probability distribution by restricting its support to $m$. Formally:

**Definition 8.** *The* $m$-bounded expectation $\mathbb{E}_{\leq m}$ *(resp.,* $\mathbb{E}_{<m}$*) of distribution* $\theta$ *with support on integers is*[11] $\mathbb{E}_{\leq m}(\theta) = \sum_{1 \leq k \leq m} \theta(k) \cdot k$ *(resp.,* $\mathbb{E}_{<m}(\theta) = \sum_{1 \leq k < m} \theta(k) \cdot k$*).*

[11]Observe that in our notation the parameter $k$ takes only (positive) integer values (although we allow the upper bound $m$ to be a non-integer). In what follows, the summation $\sum_{i=1}^{m}$ is used to denote $\sum_{i=1}^{\lfloor m \rfloor}$, where $\lfloor \cdot \rfloor$ is the greatest integer function.

**Theorem 1.** (1) *If* $\mathbb{E}_{<m_h}(\theta) \cdot \mathbb{E}_{\leq m_d}(\theta) > 1$, *then the pure stationary states are unstable, and the populations converge to an interior stationary state from any interior initial state.*[12]

(2) *If* $\mathbb{E}_{<m_h}(\theta) \cdot \mathbb{E}_{\leq m_d}(\theta) < 1$, *then the pure stationary states are asymptotically stable.*

Figure 1 illustrates Theorem 1 (the left panel illustrates Case (1) and the right panel illustrates Case (2)).

*Proof.* The proof relies on Claim 1 below, whose proof is presented in Appendix A.

**Claim 1.** $\lim_{t\to\infty} \mathbf{p}(t)$ *exists for any* $\boldsymbol{p}(0)$, *and it is a stationary state.*

It remains to show that when $\mathbb{E}_{<m_h}(\theta) \cdot \mathbb{E}_{\leq m_d}(\theta) > 1$ (resp., $\mathbb{E}_{<m_h}(\theta) \cdot \mathbb{E}_{\leq m_d}(\theta) < 1$), the pure stationary states where all agents in one population play $h$ and all agents in the other population play $d$ are unstable (resp., asymptotically stable). In what follows, we prove this.

We compute the Jacobian of the BEP dynamics for the state in which all agents of population $i$ play $d$ and all agents of population $j$ play $h$, where $i, j \in \{1, 2\}$ and $j \neq i$. For this, we consider a slightly perturbed state with a "very small" $\epsilon_i$ share of hawks in population $i$ and a "very small" $\epsilon_j$ share of doves in population $j$. By "very small," we mean that higher-order terms (terms of an order greater than or equal to two) of $\epsilon_i$ and $\epsilon_j$ are negligible.

Consider a new agent of population $i$ with a sample size of $k_i$. Action $h$ has a higher mean payoff against a sample size of $k_i$ iff (excluding rare events of having multiple $d$-s in the sample): (1) the sample includes the single action $d$ of an opponent, and (2) $k_i \leq m_h$ (by Lemma 1). The probability of (1) is $k_i \cdot \epsilon_j + o(\epsilon_j)$, where $o(\epsilon_j)$ denotes the higher-order terms of $\epsilon_j$, and so will not affect the Jacobian as $\epsilon_j \to 0$. This implies that the probability that a new agent of population $i$ (with a random sample size distributed according to $\theta$) has a higher mean payoff for action $h$ against her sample is $w_\theta(1-\epsilon_j) = \epsilon_j \cdot \sum_{k_i=1}^{m_h} \theta(k_i) \cdot k_i + o(\epsilon_j)$. An analogous argument implies that the probability that a new agent of population $j$ has a higher mean payoff for action $d_j$ against her sample is $w_\theta(\epsilon_i) = \epsilon_i \cdot \sum_{k_j=1}^{m_d} \theta(k_j) \cdot k_j + o(\epsilon_i)$. Therefore, the BEP dynamics at $(\epsilon_i, 1 - \epsilon_j)$ can be written as follows (excluding the higher-order terms of $\epsilon_i$ and $\epsilon_j$):

$$\text{(3.1)} \qquad \dot{\epsilon}_i = \epsilon_j \cdot \sum_{k_i=1}^{m_h} \theta(k_i) \cdot k_i - \epsilon_i, \qquad \dot{\epsilon}_j = \epsilon_i \cdot \sum_{k_j=1}^{m_d} \theta(k_j) \cdot k_j - \epsilon_j.$$

[12] The parameter $m_h$ is preceded by a strict inequality, while $m_d$ is preceded by a weak inequality, due to our assumption that players play $d$ if there is a tie between the actions' payoffs against their samples.

Let $a_\theta = \sum_{k_i=1}^{m_h} \theta(k_i) \cdot k_i$ and $b_\theta = \sum_{k_j=1}^{m_d} \theta(k_j) \cdot k_j$. The Jacobian of the system, Eq. (3.1), is then given by $J = \begin{pmatrix} -1 & a_\theta \\ b_\theta & -1 \end{pmatrix}$. The eigenvalues of $J$ are $-1 - \sqrt{a_\theta b_\theta}$ and $-1 + \sqrt{a_\theta b_\theta}$. Observe that: (1) if $a_\theta b_\theta > 1$ then one of the eigenvalues is positive, which implies that this state is unstable (see, e.g., Perko, 2013, Theorems 1 and 2 in Section 2.9), and (2) if $a_\theta b_\theta < 1$ then both eigenvalues are negative, which implies that the state in which all agents in population $i$ (resp., $j$) are doves (resp., hawks) is asymptotically stable. $\square$

**Corollary 1.** *For $g = l$ (the special case presented in Table 1) the population converges to an interior stationary state from any interior initial state if any of the following three conditions holds:*

1. $0 < g < \frac{1}{3}$ *and* $\theta(1) \cdot \mathbb{E}_{\leq \frac{1+g}{g}}(\theta) > 1$, *or*
2. $\frac{1}{3} \leq g < \frac{3}{5}$ *and* $(\theta(1) + 2\theta(2)) \cdot (\theta(1) + 2\theta(2) + 3\theta(3)) > 1$, *or*
3. $\frac{3}{5} \leq g < 1$ *and* $(\theta(1) + 2\theta(2)) \cdot \mathbb{E}_{\leq \frac{1+g}{1-g}}(\theta) > 1$.

*In each of these three cases, replacing ">" with "<" in the last inequality implies that the pure stationary states are asymptotically stable.*

*Proof.*

1. The inequality $g < \frac{1}{3}$ implies that $1 < m_h \equiv \frac{1+g}{1-g} < 2$. This, in turn, implies that $\mathbb{E}_{<m_h} = \theta(1)$.
2. The inequality $\frac{1}{3} \leq g < \frac{3}{5}$ implies that $2 \leq \min(m_h \equiv \frac{1+g}{1-g}, m_d \equiv \frac{1+g}{g}) \leq 3$ and $3 \leq \max(m_h, m_d) \leq 4$. This, in turn, implies either that $\mathbb{E}_{<m_h} = \theta(1) + 2\theta(2)$ and $\mathbb{E}_{\leq m_d} = \theta(1) + 2\theta(2) + 3\theta(3)$, or that $\mathbb{E}_{\leq m_d} = \theta(1) + 2\theta(2)$ and $\mathbb{E}_{<m_h} = \theta(1) + 2\theta(2) + 3\theta(3)$.
3. The inequality $\frac{3}{5} \leq g < 1$ implies that $2 < m_d \equiv \frac{1+g}{g} < 3$. This, in turn, implies that $\mathbb{E}_{\leq m_d} = \theta(1) + 2\theta(2)$. $\square$

One domain of parameters in which the populations never converge to the pure states (henceforth, global stability of miscoordination) is (1) sufficiently many agents have sample size 1, (2) the expected number of observed actions is sufficiently large, and (3) $g$ is sufficiently far from 0.5. This domain is qualitatively similar (though, larger) than the conditions for global stability of miscoordiantion under the sampling best response dynamics analyzed in Arigapudi et al. (2023).

The intuition for why these three conditions imply global stability of miscoordination is as follows. Consider an initial state in which almost everyone in population 1 plays $d$, and almost everyone in population 2 plays $h$. If $g$ is sufficiently far from 0.5, then in one of the populations a single appearance of the rare action in the opponent's population induces a new agent to play the rare action in her own population only if her sample size is small. This explains why we need Condition (1) (namely, sufficiently many agents with sample size one). By contrast, in the remaining population, a single occurrence of a rare action in the sample can induce the new agent to play the rare action in her own population also for larger sample sizes. The probability of such an occurrence of a rare action in the sample is the expected number of observed actions (assuming that $g$ is sufficiently far from 0.5), which explains why we need Condition (2).

Global stability of miscoordination holds in our setup in additional domains of parameters (which do not yield stable coordination in the best experienced sampling dynamics à la Arigapudi et al., 2023). Specifically, if $g > \frac{1}{3}$, then there is global convergence to miscoordination whenever at least half of the agents have sample size 2. The intuition for this is as follows. Consider again an initial state in which $1-\epsilon$ of the agents in population 1 play $d$, and $1-\epsilon$ of the agents in population 2 play $h$, where $0 < \epsilon << 1$. If $g > \frac{1}{3}$, then in both populations a single occurrence of the opponent's rare action in one of the samples can induce a new agent with sample size two to play the rare action in her own population. The probability of having such a rare occurrence in the agent's sample (of size two) is approximately $2\epsilon$. Thus, at least $2\theta(2)\epsilon$ of the new agents in each population will play the rare action. This, in turn, implies that the share of agents playing the rare action increases if $\theta(2) > 0.5$.

## 3.2 Asymptotic Stability of the Inefficient Symmetric State

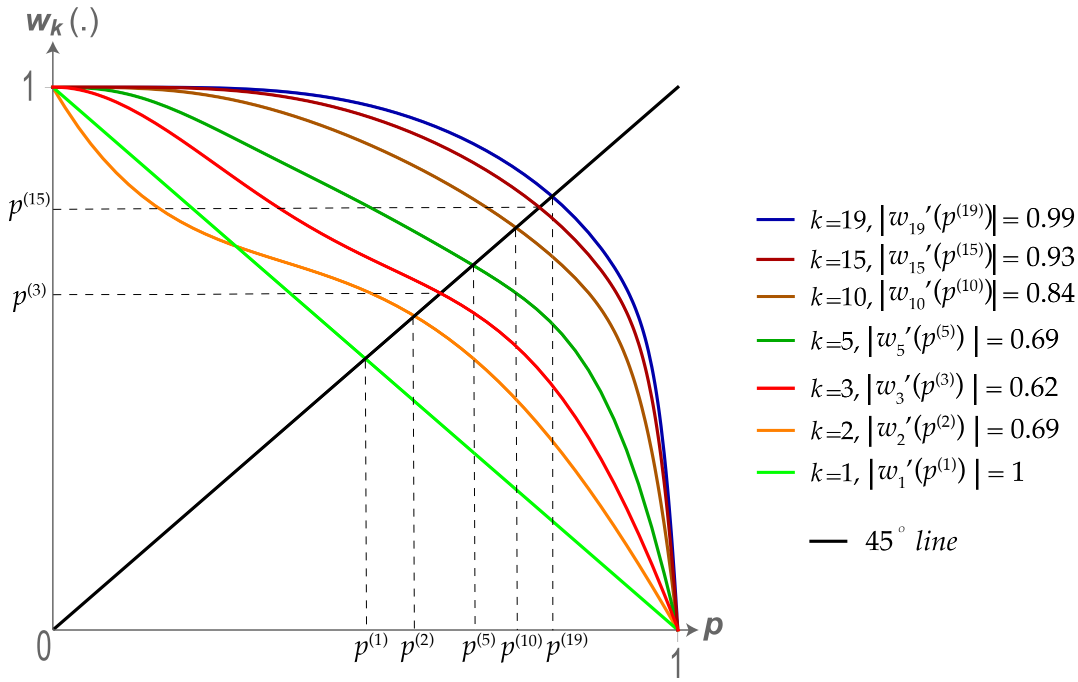
We show that the symmetric stationary state is asymptotically stable under the BEP dynamics for homogeneous populations (and populations in which all agents have relatively small sample sizes). For tractability, we focus on cases where the gain of a hawkish player against her dovish opponent is large, namely, $l, g > 1 - \frac{1}{\max(\text{supp}(\theta))}$. Our results shows that in this domain, the symmetric stationary state is asymptotically stable in various populations in which agents have relatively small samples:

1. for any homogeneous distribution of sample sizes $\theta \equiv k < 20$, or

2. for any heterogeneous distribution with a maximal sample size of 5.

The threshold of $k = 20$ is binding. The symmetric stationary state becomes unstable if the sample size is $k \geq 20$. By contrast, the bound of a maximal size of 5 for heterogeneous distributions of sample sizes is only a constraint of our proof technique.

Figure 2: The Function $w_k(p)$ for Various Values of $k$




**Theorem 2.** *Assume that* $l, g > 1 - \frac{1}{\max(supp(\theta))}$*, and either (1)* $\theta \equiv k < 20$*, or (2)* $\max(\text{supp}(\theta)) \leq 5$. *Then, the game admits an asymptotically stable symmetric state under the BEP dynamics.*

*Sketch of Proof.* When $l$ and $g$ are sufficiently large, the payoff of action $h$ is slightly below twice the number of $d$-s in the $h$-sample, and the payoff of action $d$ is slightly above the number of $d$-s in the $d$-sample. This implies that action $h$ has a higher mean payoff than action $d$ iff the number of $d$-s in the $h$-sample is strictly greater than half the number of $d$-s in the $d$-sample.

Thus, we can write $w_k(p)$ as follows:

$$w_k(p) = P\left(\underbrace{k - X(k,p)}_{\#d\text{-}s \text{ in } h\text{-sample}} > \frac{1}{2}\underbrace{(k - Y(k,p))}_{\#d\text{-}s \text{ in } d\text{-sample}}\right) = P(2X(k,p) - Y(k,p) < k), \tag{3.2}$$

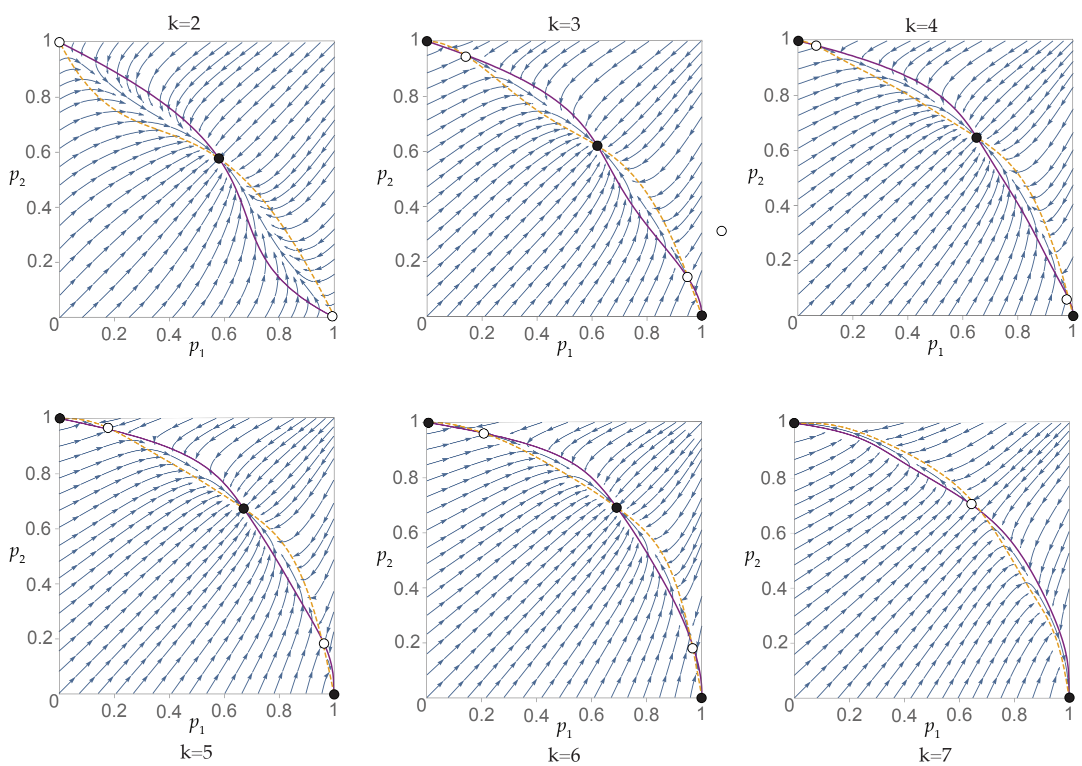
where $X(k,p)$ and $Y(k,p)$ are iid binomial random variables with parameters $k$ and $p$.

In the formal proof (see Appendix B), we show that for any $k < 20$, $w_k(p)$ has a unique fixed point $p^{(k)}$ and that $|w_k'(p^{(k)})| < 1$ (see Figure 2). This implies, by the same argument as in the proof of Theorem 1, that the symmetric stationary state is asymptotically stable. (By contrast, one can verify that $|w_k'(p^{(k)})| > 1$ for $k \geq 20$, which implies that the symmetric stationary state is unstable for $k \geq 20$.)

Next, we verify in the formal proof that for any $k \in \{1, 2, 3, 4, 5\}$ it holds that (I) the fixed points are all in the interval $(0.5, 0.68)$, and (II) $|w'_k(p)| < 1$ for any $k \in \{1, ..., 5\}$ and any $p \in (0.5, 0.68)$. Let $\theta$ be any distribution with $\max(\text{supp}(\theta)) \leq 5$. The fact that $w_\theta(p)$ is a weighted average of the various $w_k(p)$ implies that (I) the fixed point $p^{(\theta)}$ of $w_\theta(p)$ is in $(0.5, 0.68)$, and (II) $|w'_\theta(p^{(\theta)})| < 1 \Rightarrow (p^{(\theta)}, p^{(\theta)})$ is asymptotically stable. □

*Example 1.* Consider the hawk–dove game in which $g = l = 0.85$, and all agents have the same sample size $k$. Observe that the inequality condition of Theorem 2 (namely, $l, g > 1 - \frac{1}{\max(\text{supp}(\theta))} = 1 - \frac{1}{k}$) holds iff $k < 1/0.15 \approx 6.67$. Figure 3 shows the phase plots for the values of $k$ between 2 and 7. Observe that for each $1 < k < 7$, the symmetric interior state is asymptotically stable, and, moreover, its basin of attraction is very large (a Monte-Carlo analysis shows that the basin of attraction in all cases is at least $96\%$). By contrast, when $k = 7$ (and for all larger values of $k$) we get global convergence to the pure states.

Figure 3: Phase Plots for Various Values of $k$ ($g = l = 0.85$)




The figure illustrates Theorem 2 by showing the phase plots of BEP dynamics for $g = l = 0.85$ and various values of $\theta \equiv k$.

### 3.3 Numeric Analysis of Geometric Distributions

In this subsection we demonstrate the applicability of the conditions for stability of miscoordiantion by studying the family of geometric distributions of sample sizes.

We say that the distribution of sample sizes is geometrically distributed with parameter $q \in (0,1)$ if the probability of obtaining a sample size of $k$ is $\theta(k) = (1-q)^{k-1} \cdot q$. One interpretation of this distribution is that each new agent tests her feasible actions sequentially and, after each round of sampling (in which she samples both actions), the agent stops with probability $q$, or continues to another round of sampling with the remaining probability of $1-q$.

We run 1,000 Monte-Carlo simulations for each combination of values of

$$g, q \in \{0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, 0.95\}$$

to estimate the basin of attraction of the symmetric interior state and the pure stationary states. For the simulations, we bounded each geometric distribution by a maximal value of $\bar{k} = 20$ (i.e., we set $\theta(\bar{k}) = 1 - \sum_{k<\bar{k}} \theta(k)$, and $\theta(k) = 0 \ \forall k > \bar{k}$). The results (see Table 3) show that in $58\%$ of the cases there is global (or almost global) convergence to the symmetric interior state. This occurs whenever either $g$ (the gain of a hawkish player against a dove) or $q$ (the probability of a new agent stopping her sample after each round) is sufficiently large. In $5\%$ of the cases the symmetric interior state is asymptotically stable with a large basin of attraction ($\geq 69\%$). Finally, in the remaining $37\%$ of the cases in which $g$ and $q$ are sufficiently small, the symmetric interior state is unstable. In almost all cases, except for two cases marked with an asterisk $(^{*})$ in the table, the populations converge to one of the pure stationary states. In the two cases of $g = 0.35, q = 0.35$ and $g = 0.05, q = 0.95$, there is global convergence to one of the asymmetric interior stationary states.

# 4. Conclusion

A key result in evolutionary game theory is that two populations that are matched to play a hawk–dove game converge to one of the pure equilibria from almost any initial state. We demonstrate that this result crucially depends on the evolutionary dynamics being monotone. Specifically, we show that plausible classes of dynamics in which agents occasionally revise their actions based on the payoffs they obtained in a few past interactions, can lead to the opposite prediction: convergence to an inefficient interior stationary state. Moreover, our analysis suggests that this occurs under a large domain of parameter

Table 3: Size of Basin of Attraction of Symmetric Interior Stationary State

| $g \backslash q$ | 0.05 | 0.15 | 0.25 | 0.35 | 0.45 | 0.55 | 0.65 | 0.75 | 0.85 | 0.95 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0* |
| 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.77 | 0.99 | 1 |
| 0.25 | 0 | 0 | 0 | 0 | 0 | 0 | 0.69 | 0.91 | 0.98 | 1 |
| 0.35 | 0 | 0 | 0 | 0* | 1 | 1 | 1 | 1 | 1 | 1 |
| 0.45 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0.55 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0.65 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0.75 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0.85 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0.95 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Initial states outside the basin of attraction of the symmetric interior stationary state converge to one of the pure states, except in the two cases marked with *. In those two cases there is global convergence to asymmetric interior states.

values: when either the sample sizes are sufficiently small or the gain of a hawkish player against a dovish opponent is sufficiently large.

A key finding of this paper is to show that in a two-population hawk–dove game, interior stationary states can be asymptotically stable under BEP dynamics even when agents in the population have homogeneous sample sizes. This is in contrast to the findings of Arigapudi et al. (2023) who show that in a two-population coordination game, interior stationary states can be asymptotically stable under the sampling best response dynamics only if there is heterogeneity in the sample sizes. As discussed at the end of the Introduction, the predictions of our model can be tested in the lab in the simple "black box" experimental setting, in which players do not know the game's structure and observe only their realized payoffs.

Our results provide a new explanation of why in bargaining situations, such as the motivating example of buying and selling houses, players in both populations tend to play hawkish strategies, and bargaining frequently fails. Our model assumes that all players in each population have the same payoff matrix. Heterogeneity in the payoffs, and private information regarding one's payoff, are important aspects in many real-life bargaining situations. An interesting direction for future research is to apply the analysis of BEP dynamics in richer models that incorporate heterogeneous payoffs.

# Appendix

## A. Proof of Theorem 1

### A.1 Proof of Claim 1

We define the following sets:

$$\mathcal{I} = \{(p_1, p_2) \in [0,1]^2 : w(p_1) \leq p_2 \leq w^{-1}(p_1) \text{ or } w^{-1}(p_1) \leq p_2 \leq w(p_1)\}$$
$$\mathcal{E} = \{(p_1, p_2) \in [0,1]^2 : p_2 = w(p_1) = w^{-1}(p_1)\}.$$

$\mathcal{I}$ is the region between the graphs $w$ and $w^{-1}$. $\mathcal{E}$ is the set that contains the stationary states. Further, notice that $\mathcal{I} \supset \mathcal{E}$.

Let $y, z \in \mathbb{R}^2$ and $Z \subset \mathbb{R}^2$. The Euclidean distances are defined as follows:

$$d(y, z) = ||y - z||_2 \text{ and } d(y, Z) = \inf_{z \in Z} d(w, z).$$

To complete the proof, we must show that $d((p_1(t), p_2(t)), \mathcal{E}) \to 0$ as $t \to \infty$, where $(p_1(t), p_2(t))_{t \geq 0}$ is the solution to the differential equations $\dot{p}_1 = w(p_2) - p_1$ and $\dot{p}_2 = w(p_1) - p_2$. In what follows, we show this.

We define $T$ be the hitting time to hit the set $\mathcal{I}$, i.e., $T := \inf\{t \geq 0 : (x(t), y(t)) \in \mathcal{I}\}$. We consider the following two exhaustive cases.

**Case 1:** $T = \infty$. In this case, we define two functions as follows: $L_1(t) = \frac{1}{2}(p_1(t) - w(p_2(t)))^2, L_2(t) = \frac{1}{2}(p_1(t) - w^{-1}(p_2(t)))^2$. Assume that $(p_1(0), p_2(0))$ is in the lower triangular part of the phase space, i.e., $p_2(0) < w(p_1(0))$ and $p_2(0) < w^{-1}(p_1(0))$. Everything that follows applies similarly if we assume an initial condition in the upper triangular part. For $t < T$, we have

$$\begin{aligned} L_1'(t) &= (p_1(t) - w(p_2(t)))(\dot{p}_1(t) - w'(p_2(t))\dot{p}_2(t)) \\ &= (p_1(t) - w(p_2(t)))(w(p_2(t)) - p_1(t) - w'(p_2(t))(w(p_1(t)) - p_2(t))) \\ &= -(w(p_2(t)) - p_1(t))^2 + w'(p_2(t))(w(p_2(t)) - p_1(t))(w(p_1(t)) - p_2(t)) \\ &\leq -(p_1(t) - w(p_2(t)))^2 = -2L_1(t). \end{aligned}$$

From Gronwall's inequality, it follows that $L_1(t) \leq L_1(0)e^{-2t}$. As $T = \infty$, we have $L_1(t) \to 0$ as $t \to \infty$. Similarly, we can show that $L_2(t) \to 0$ as $t \to \infty$ for $T = \infty$.

From the above, we can conclude that $d((p_1(t), p_2(t)), \mathcal{E}) \to 0$ as $t \to \infty$.

**Case 2:** $T < \infty$. We have $(p_1(T), p_2(T)) \in \mathcal{I}$. Let $w^{-1}(p_1(T)) \leq p_2(T) \leq w(p_1(T))$. The other case $w(p_1(T)) \leq p_2(T) \leq w^{-1}(p_1(T))$ can be dealt with similarly. If $w^{-1}(p_1(T)) = p_2(T) = w(p_1(T))$ then $(p_1(T), p_2(T))$ is already in equilibrium. Assume that $w^{-1}(p_1(T)) \leq p_2(T) < w(p_1(T))$. Notice that $w^{-1}(p_1(t)) \leq p_2(t) \leq w(p_1(t))$ for all $t \geq T$. Denote this invariant set by $\mathcal{I}_1$, i.e., $\mathcal{I}_1 = \{(p_1, p_2) \in [0,1]^2 : w^{-1}(p_1) \leq p_2 \leq w(p_1)\}$. Further, let $(p_1^*, p_2^*)$ be the equilibrium that lies in the left part of $\mathcal{I}_1$ i.e., $p_1^* = \sup_{(p_1,p_2) \in \mathcal{E}} p_1 < p_1(T)$. It is straightforward to verify that the function $L(p_1, p_2) = \frac{1}{2}((p_1 - p_1^*)^2 + (p_2 - p_2^*)^2)$ is a strict

Lyapunov function on $\mathcal{I}_1$ which implies that in this case the solution converges to $(p_1^*, p_2^*)$. Similarly, we can show that if $w(p_1(T)) \leq p_2(T) < w^{-1}(p_1(T))$, then there is convergence to the rightmost closest equilibrium of $\mathcal{I}_2 = \{(p_1, p_2) \in [0,1]^2 : w(p_1) \leq p_2 \leq w^{-1}(p_1)\}$.

From the above, we can once again conclude that $d((p_1(t), p_2(t)), \mathcal{E}) \to 0$ as $t \to \infty$, which completes the proof.

# B. Proof of Theorem 2

Recall that the BEP dynamics are given by $\dot{p}_1 = \delta(w_\theta(p_2) - p_1)$ and $\dot{p}_2 = \delta(w_\theta(p_1) - p_2)$, and that a symmetric state $(p^{(\theta)}, p^{(\theta)})$ is stationary iff $w_\theta(p^{(\theta)}) = p^{(\theta)}$.

We now establish some properties of the BEP dynamics and the symmetric rest points for symmetric distributions of types $\theta \equiv k$. If $l, g \in \left(1 - \frac{1}{\max(\text{supp}(\theta))}, 1\right)$, then action $h_i$ has a higher mean payoff iff the number of $d_j$-s in the $h_i$-sample is strictly greater than half the number of $d_j$-s in the $d_i$-sample. To express $w_k(p)$ concisely in this case, we let $X(k,p)$ and $Y(k,p)$ denote iid binomial random variables with parameters $k$ and $p$. We can then write $w_k(p)$ as follows:

$$\text{(B.1)} \qquad w_k(p) = P\left(k - X(k,p) > \frac{1}{2}(k - Y(k,p))\right) = P(2X(k,p) - Y(k,p) < k).$$

Observe that $w_k(p)$ is a polynomial in $p$ of degree at most $2 \cdot k$. We have verified the following facts about these polynomials for $k < 20$ (for an illustration see Figure 2; the Mathematica code is given in the online supplementary material, and the explicit values of the rest points and the derivatives are presented in Table 4):

- For $k \in \{1, 2, \ldots, 18, 19\}$, $w_k(p)$ is decreasing in $p$.
- For $k \in \{1, 2, \ldots, 18, 19\}$, $w_k(p)$ has a unique fixed point $p^{(k)}$. Moreover, $0.5 < p^{(k)} < 0.68$ for any $k \in \{1, 2, 3, 4, 5\}$.
- $|w_1'(p)| \equiv 1$, and $|w_k'(p)| < 1$ for any $k \in \{2, 3, 4, 5\}$ and $0.5 < p < 0.68$.

Recall that $w_\theta(p)$ is a convex combination of the $w_k(p)$ for the $k$-s in its support (i.e., $w_\theta(p) = \sum_k \theta(k) \cdot w_k(p)$). From the above facts, it follows that:

1. For $\theta \equiv k < 20$, the function $w_k(p)$ has a unique fixed point $p^{(k)}$ such that $|w_k'(p^{(k)})| < 1$, which implies that $(p^{(k)}, p^{(k)})$ is asymptotically stable.
2. For $\max(\text{supp}(\theta)) \leq 5$, the function $w_\theta(p)$ has a unique fixed point $p^{(\theta)}$ such that $p^{(\theta)} \in (0.5, 0.68)$ and $|w_\theta'(p^{(\theta)})| < 1$ if $\theta(1) \neq 1$, which implies that $(p^{(\theta)}, p^{(\theta)})$ is asymptotically stable.

Table 4: Fixed Points of the Function $w_k(p)$ in the Proof of Theorem 2

| $k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $p^{(k)}$ | 0.500 | 0.579 | 0.620 | 0.649 | 0.672 | 0.690 | 0.706 | 0.720 | 0.731 | 0.741 |
| $\lvert w_k'(p^{(k)})\rvert$ | 1 | 0.690 | 0.618 | 0.645 | 0.690 | 0.730 | 0.763 | 0.793 | 0.818 | 0.840 |

| $k$ | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|
| $p^{(k)}$ | 0.750 | 0.758 | 0.765 | 0.773 | 0.778 | 0.784 | 0.789 | 0.794 | 0.799 | 0.803 |
| $\lvert w_k'(p^{(k)})\rvert$ | 0.861 | 0.88 | 0.899 | 0.916 | 0.932 | 0.948 | 0.963 | 0.978 | 0.991 | 1.001 |

Table 5: Values of $|w_k'(p^{(j)})|$ for $k, j \in \{1, 2, 3, 4, 5\}$.

| $k \backslash p^{(j)}$ | $p^{(1)}$ | $p^{(2)}$ | $p^{(3)}$ | $p^{(4)}$ | $p^{(5)}$ |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 0.5 | 0.690 | 0.812 | 0.905 | 0.981 |
| 3 | 0.562 | 0.560 | 0.618 | 0.687 | 0.759 |
| 4 | 0.625 | 0.616 | 0.623 | 0.645 | 0.679 |
| 5 | 0.605 | 0.642 | 0.659 | 0.673 | 0.690 |

# References


Arigapudi, S., Heller, Y., and Milchtaich, I. (2021). Instability of defection in the prisoner's dilemma under best experienced payoff dynamics. *Journal of Economic Theory*, 197:105174.

Arigapudi, S., Heller, Y., and Schreiber, A. (2023). Heterogeneous noise and stable miscoordination. *arXiv preprint arXiv:2305.10301*.

Arigapudi, S. and Lahkar, R. (2024). Emergence of trust in the trust game under best experienced payoff dynamics. *Economics Letters*, 236:111620.

Aumann, R. J. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, 55(1):1–18.

Baliga, S. and Sjöström, T. (2012). The strategy of manipulating conflict. *American Economic Review*, 102(6):2897–2922.

Baliga, S. and Sjöström, T. (2020). The strategy and technology of conflict. *Journal of Political Economy*, 128(8):3186–3219.

Berkemer, R., Starke, J., and Kawamoto, A. (2023). Sampling dynamics applied to the traveler's dilemma reveals non-nash behavior. *Games and Economic Behavior*, 142:339–356.

Bornstein, G., Budescu, D., and Zamir, S. (1997). Cooperation in intergroup, $n$-person, and two-person games of chicken. *Journal of Conflict Resolution*, 41(3):384–406.

Brams, S. J. and Kilgour, D. M. (1987). Threat escalation and crisis stability: A game-theoretic analysis. *American Political Science Review*, 81(3):833–850.

Brams, S. J. and Kilgour, D. M. (2001). Fallback bargaining. *Group Decision and Negotiation*, 10(4):287–316.

Burton-Chellew, M. N., El Mouden, C., and West, S. A. (2017). Social learning and the demise of costly cooperation in humans. *Proceedings of the Royal Society B: Biological Sciences*, 284:20170067.

Cárdenas, J., Mantilla, C., and Sethi, R. (2015). Stable sampling equilibrium in common pool resource games. *Games*, 6(3):299–317.

Danenberg, T. and Spiegler, R. (2022). Representative sampling equilibrium. Working Paper.

Dixit, A., McAdams, D., and Skeath, S. (2019). "We haven't got but one more day": The Cuban missile crisis as a dynamic chicken game. *Available at SSRN 3406265*.

Herold, F. and Kuzmics, C. (2020). The evolution of taking roles. *Journal of Economic Behavior & Organization*, 174:38–63.

Izquierdo, S. S. and Izquierdo, L. R. (2022). Stability of strict equilibria in best experienced payoff dynamics: Simple formulas and applications. *Journal of Economic Theory*, 206:105553.

Izquierdo, S. S. and Izquierdo, L. R. (2023). Strategy sets closed under payoff sampling. *Games and Economic Behavior*, 138:126–142.

Koster, H. R. and Rouwendal, J. (2023). Housing market discount rates: Evidence from bargaining and bidding wars. *International Economic Review*.

Larsen, B. J. (2021). The efficiency of real-world bargaining: Evidence from wholesale used-auto auctions. *Review of Economic Studies*, 88(2):851–882.

Lipnowski, I. and Maital, S. (1983). Voluntary provision of a pure public good as the game of "chicken". *Journal of Public Economics*, 20(3):381–386.

Mantilla, C., Sethi, R., and Cárdenas, J. C. (2018). Efficiency and stability of sampling equilibrium in public goods games. *Journal of Public Economic Theory*, 22(2):355–370.

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press.

Maynard Smith, J. and Parker, G. A. (1976). The logic of asymmetric contests. *Animal Behaviour*, 24(1):159–175.

Nax, H. H., Burton-Chellew, M. N., West, S. A., and Young, H. P. (2016). Learning in a black box. *Journal of Economic Behavior & Organization*, 127:1–15.

Nax, H. H. and Perc, M. (2015). Directional learning and the provisioning of public goods. *Scientific Reports*, 5(1):1–6.

Oprea, R., Henwood, K., and Friedman, D. (2011). Separating the hawks from the doves: Evidence from continuous time laboratory games. *Journal of Economic Theory*, 146(6):2206–2225.

Osborne, M. J. and Rubinstein, A. (1998). Games with procedurally rational players. *American Economic Review*, 88(4):834–847.

Osborne, M. J. and Rubinstein, A. (2003). Sampling equilibrium, with an application to strategic voting. *Games and Economic Behavior*, 45(2):434–441.

Oyama, D., Sandholm, W. H., and Tercieux, O. (2015). Sampling best response dynamics and deterministic equilibrium selection. *Theoretical Economics*, 10(1):243–281.

Perko, L. (2013). *Differential Equations and Dynamical Systems*, volume 7. Springer Science & Business Media.

Rapoport, A. and Chammah, A. M. (1966). The game of chicken. *American Behavioral Scientist*, 10(3):10–28.

Salant, Y. and Cherry, J. (2020). Statistical inference in games. *Econometrica*, 88(4):1725–1752.

Sandholm, W. H. (2001). Almost global convergence to $p$-dominant equilibrium. *International Journal of Game Theory*, 30(1):107–116.

Sandholm, W. H. (2010). *Population Games and Evolutionary Dynamics*. MIT Press.

Sandholm, W. H., Izquierdo, S. S., and Izquierdo, L. R. (2019). Best experienced payoff dynamics and cooperation in the centipede game. *Theoretical Economics*, 14(4):1347–1385.

Sandholm, W. H., Izquierdo, S. S., and Izquierdo, L. R. (2020). Stability for best experienced payoff dynamics. *Journal of Economic Theory*, 185:104957.

Sawa, R. and Wu, J. (2023). Statistical inference in evolutionary dynamics. *Games and Economic Behavior*, 137:294–316.

Sethi, R. (2000). Stability of equilibria in games with procedurally rational players. *Games and Economic Behavior*, 32(1):85–104.

Sethi, R. (2021). Stable sampling in repeated games. *Journal of Economic Theory*, 197:105343.

Spiegler, R. (2006). Competition over agents with boundedly rational expectations. *Theoretical Economics*, 1(2):207–231.

Sugden, R. (1989). Spontaneous order. *Journal of Economic Perspectives*, 3(4):85–97.

Taylor, P. D. (1979). Evolutionarily stable strategies with two types of player. *Journal of Applied Probability*, 16:76–83.

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