A Justification for Equal Sharing

Jiabin Wu[*] Hanzhe Zhang[†]

March 9, 2026

Abstract

We study the evolution of productive traits under different surplus-sharing norms in one-sided matching markets with heterogeneous agents. Agents form partnerships endogenously, and the output of each partnership depends on the productivity of the matched agents. The division of output is governed by a social norm that specifies how surplus is shared between partners.

We compare three sharing rules: a market sharing rule that implements transferable-utility matching, an equal sharing rule, and unequal sharing norms. Although equal sharing may generate inefficient matchings in the short run, we show that it promotes the evolutionary growth of productive traits and thereby achieves long-run efficiency comparable to that of market sharing. By contrast, unequal sharing norms generate inefficiencies both in matching and in evolutionary outcomes.

Our results provide an evolutionary rationale for the persistence of equal sharing norms alongside market-based institutions.

Keywords: Equal sharing, norms, matching, evolutionary game theory

JEL Codes: C73, C78, Z10

---

[*] University of Oregon, jwu5@uoregon.edu.

[†] Michigan State University, hanzhe@msu.edu.

1 Introduction

The equal sharing norm advocates for the equitable distribution of resources among individuals. It has deep roots in human history and continues to influence contemporary societal structures. For example, the ancient principle of equal sharing in partnership law, as documented in the *Institutes of Justinian* since 533 CE, remains widely adopted by modern law firms (Smith, 2001). Anthropological evidence suggests that egalitarian sharing norms were also central to many hunter–gatherer societies, where food obtained through hunting or gathering was typically distributed broadly within the camp rather than retained by the individual who produced it (Lee, 1979; Woodburn, 1982). Consistent with these observations, experimental evidence from fifteen small-scale societies shows that individuals often display strong fairness and sharing preferences rather than purely self-interested behavior (Henrich et al., 2001). Equal sharing of loyalties among band members is also a common practice, even among the most renowned bands (Polcz, 2023). Finally, the kibbutz has been a vibrant modern Israeli community that promotes equality in resource allocation among community members (Abramitzky, 2008, 2011, 2018).

While equal sharing offers appealing benefits, such as fostering fairness and reinforcing group cohesion, it is inherently rigid. Market-based approaches, which allow for compensation to vary based on different factors, are often seen as more efficient and adaptable to changing conditions. Despite this, equal sharing remains a common practice in certain areas of the market economy. This raises the question: what underlying mechanisms sustain the persistence of equal sharing, even when more flexible alternatives are available? In this paper, we present an evolutionary explanation.

We model a heterogeneous population of agents who endogenously form partnerships. The output of each partnership depends on the productivity of the matched agents. The division of this output is governed by societal sharing norms.

We analyze three types of sharing norms. First, the ***market sharing norm*** allows agents to split output based on their market opportunities, modeled as a perfectly transferable utility (TU) one-sided matching market. Second, the ***equal sharing norm*** requires matched pairs to divide output equally. Third, we consider ***unequal sharing norms***, such as sharing rules that allocate a larger (or smaller) share to the higher ranked agent in the social hierarchy. Unlike the market sharing norm, equal and unequal sharing norms lack flexibility, causing the market to operate as a nontransferable utility (NTU) one-sided matching market.

In our model, evolution operates at two levels: agents' matching behavior and the distribution of productive traits in the population. At the behavioral level, agents within a generation engage in a repeated re-matching process, which immediately results in a stable matching. At the productivity level, the distribution of productive traits evolves more slowly across generations over time according to the replicator dynamic, which is driven by the payoffs earned through matching, and converges to a stable distribution in the long run.

We find that although the equal sharing norm may yield stable matchings that are less efficient than those under the market sharing norm at the behavioral level, both norms still lead to efficiency at the level of trait evolution. The reason is that the market sharing rule attains efficiency through flexibility (by compensating individuals appropriately), whereas equal sharing attains it through robustness (by matching individuals who jointly produce the most to the greatest extent possible). As a result, both ultimately favor the selection of the distribution of productive traits that generates the highest partnership outputs.

In contrast, unequal sharing norms can introduce inefficiencies in the process of trait evolution, as traits conducive to productivity may be suppressed when they are disadvantaged within a hierarchical society. We show that unequal sharing norms result in production inefficiency in the matching stage and less productive traits survive the evolutionary process. This suggests that, over time, unequal sharing norms are largely outperformed by both the market sharing norm and the equal sharing norm, which will prevail in the long run.

To our knowledge, this paper is the first to study how different surplus-sharing norms shape long-run evolutionary outcomes in a matching market with heterogeneous productivity. Our paper connects three strands of literature. First, it relates to the theory of evolutionary dynamics in population games. A large literature studies how strategy distributions evolve under payoffmonotonic dynamics such as the replicator dynamic, where strategies with above-average payoffs increase in frequency over time. Sandholm (2010) provides a unified treatment of these models and their stability properties in large populations. Second, our work is related to the evolutionary study of social norms governing surplus division. Skyrms (2014) analyzes the evolution of division rules in bargaining environments with random matching. In contrast, our framework features heterogeneous agents whose partnerships form endogenously through a matching process, which fundamentally shapes the evolutionary selection of traits. Finally, our model connects evolutionary dynamics with the theory of matching markets. Classic results in transferable-utility matching, beginning with Becker (1973) and the assignment game of Shapley and Shubik (1972), characterize stable and efficient allocations in matching environments. We embed this matching structure into an evolutionary framework and show how alternative sharing norms interact with matching behavior to determine long-run population outcomes.

Previous studies adopt a similar approach to studying the evolution of cultural or preference traits under rapid matching and a fixed sharing rule (e.g., Wu and Zhang, 2021; Hiller, Wu, and Zhang, 2023; Wang and Wu, 2025). However, they do not investigate how different sharing rules shape long-run outcomes differently.

In our model, partnership formation occurs in a one-sided pairwise matching market (i.e., the roommate market) with a continuum of agents in every period. Although many standard results from the parallel two-sided matching settings—e.g., equivalence of socially efficient and stable matchings, convexity of stable payoffs (Shapley and Shubik, 1972; Gretsky, Ostroy, and Zame, 1992)—extend to our one-sided setting, there is no paper that explicitly states and proves these results.[1] To this end, we provide results in more detail and prove them rigorously, which may serve as a reference for subsequent papers that model the evolution of behavior, productivity, culture, and/or preferences in one-sided matching markets. In addition, it should be noted that although we explicitly model the formation of two-member teams, the main result regarding the long-run effi-

---

[1] A few papers—Azevedo, Weyl, and White (2013); Chiappori, Salanié, and Galichon (2019); Carmona and Laohakunakorn (2024)—provide existence results in similar roommate matching settings but do not discuss properties such as convexity and compactness of the set of stable payoffs, which are crucial to our evolutionary existence results.

ciency of the equal sharing norm would hold if we considered the formation of multiple-member teams.

The only other work we are aware of that provides evolutionary justifications for the equal sharing norm is the book of Skyrms (2014). In his model, a population of agents is randomly matched to play a dollar division game. The population consists of three types of agents: super-greedy agents who demand at least 2/3, fair-minded agents who demand at least 1/2, and super-modest agents who demand at least 1/3. The author demonstrates that, under the replicator dynamic, a homogeneous population of fair-minded agents as well as a polymorphic population with equal proportions of super-greedy and super-modest agents are both asymptotically stable. However, the basin of attraction for the fair-minded population is significantly larger than that of the polymorphic population. The intuition is straightforward. Super-greedy agents secure 2/3 only when paired with super-modest agents and receive nothing otherwise. Super-modest agents can always reach an agreement with any opponent, but they consistently receive the smallest share. Fair-minded agents, by contrast, typically secure 1/2, especially when there is a sufficient number of fair-minded or super-modest agents in the population.

Our model differs from Skyrms (2014) in several key aspects. First, agents in our model exhibit heterogeneity in their productivity, rather than heterogeneity in demand rules. Second, partnerships in our model emerge endogenously, as opposed to being formed through random encounters, and this is key to the long-run efficiency of the equal sharing norm in our evolutionary setting.

The rest of the paper is organized as follows. Section 2 studies the market sharing norm. Section 3 studies alternative fixed-proportion sharing norms, i.e., equal and unequal sharing norms. Section 4 concludes. The appendix collects omitted proofs.

2 The Market Sharing Norm

2.1 Population and Matching

Consider a unit mass of agents. Each agent is characterized by a productive trait $i \in N = \{1, 2, ..., n\}$. Let $x_i \in [0, 1]$ denote the mass of trait-i agents, for any $i \in N$, and $\sum_{i=1}^{n} x_i = 1$. Let $x = (x_1, x_2, ..., x_n)$ denote the population state. In this paper, we use $\mathbb{R}_+$ to refer to the set of nonnegative real numbers and $\mathbb{R}_{++}$ to refer to the set of strictly positive real numbers. Let $X = \{x \in \mathbb{R}^n_+ | \sum_{i=1}^{n} x_i = 1\}$ denote the set of population states; it is an $(n-1)$-simplex Δ^{n-1}. Let $X_{++} = \text{int}(X) = \{x \in \mathbb{R}^n_{++} | \sum_{i=1}^{n} x_i = 1\}$ denote the ***interior*** of X, i.e., the set of population states such that there is a positive mass of each trait.

Agents can find partners to form pairs; they can also choose not to: If an agent remains unmatched, she receives a payoff of 0. If a trait-i agent and a trait-j agent are matched, they jointly produce a strictly positive payoff of $f(i, j)$. Let $f : N^2 \to \mathbb{R}_{++}$ denote the production function. Assume that f is symmetric: $f(i, j) = f(j, i)$ for any $i, j \in N$.

Let $\mu_{ij}(x) \geq 0$ denote the mass of (i, j) pairs given the population state x. We have $\mu_{ij}(x) = \mu_{ji}(x)$. Since each (i, i) pair contains two trait-i agents, and each (i, j) pair with $i \neq j$ contains one trait-i agent, we have $2\mu_{ii}(x) + \sum_{j \neq i} \mu_{ij}(x) \leq x_i$. Let $\mu(x)$ be an $n \times n$ matrix that denotes the aggregate matching (Echenique et al., 2013), whose (i, j) entry is $\mu_{ij}(x)$. Note that the notion of aggregate matching is defined on the population level as opposed to the more familiar individual level. We adopt the stability concept in two-sided TU (transferable utility) matching market (Shapley and Shubik, 1972; Becker, 1973; Gretsky, Ostroy, and Zame, 1992) to our setting of one-sided matching market with a continuum of agents to characterize the stable outcome of the partnership formation process.

Definition 1 *A **stable outcome** is described by a **stable matching** $\mu^*(x)$ and a vector of **stable payoffs** $w^*(x) = (w^*_1(x), w^*_2(x), ..., w^*_n(x)) \in \mathbb{R}^n$, such that*

*(i) (**individual rationality**) each person gets at least as much as staying unmatched: $w^*_i(x) \geq 0$ for any $i \in N$;*

(ii) *(**surplus efficiency**) each pair exactly divides the surplus:* $w^*_i(x) + w^*_j(x) = f(i,j)$ *if* $\mu^*_{ij}(x) > 0$, *for any* $i, j \in N$*; and*

(iii) *(**no blocking pair**) no pair can get a sum that is strictly more than the sum of their current payoffs:* $w^*_i(x) + w^*_j(x) \geq f(i,j)$ *for any* $i, j \in N$*.*

Let us note some implications of the definition. First, since an agent gets a payoff of zero if she remains unmatched and would produce a strictly positive surplus with any other agent, all agents are matched in any stable matching. Second, in the definition, all agents with the same trait have the same stable payoff. This is not an assumption, but would be implied by the definition, namely, by the no-blocking-pair condition: If two agents with the same trait have different payoffs, then there is room for a blocking pair to emerge.

Lemma 1 *A stable outcome exists.*

We show in the proof of Lemma 1 that just as in a two-sided matching market (Gretsky, Ostroy, and Zame, 1992), a matching is stable if and only if it is socially efficient. Hence, we can characterize a stable matching by the solution to the maximization problem of the total payoff:

$$\max_{\mu(x)} \sum_{i \in N} \sum_{j \leq i} \mu_{ij}(x) f(i,j) \text{ such that } 2\mu_{ii}(x) + \sum_{j \neq i} \mu_{ij}(x) \leq x_i \text{ for all } i \in N.$$

When the solution to the maximization problem is unique, it must be the unique stable matching. We will assume that the stable matching is always unique in the subsequent analysis.

We make the following additional assumptions on the primitives of the model to eliminate knife-edge cases and simplify the analysis.

Assumption 1

(i) $f(i,i) \neq f(j,j)$ *for any* $i \neq j$*.*

(ii) $f(i,j) + f(i',j') \neq f(i,i') + f(j,j')$ *for any* $i, i', j, j' \in N$ *such that* $\{i,i'\} \cap \{j,j'\} = \emptyset$*.*

Assumption 1(i) ensures that some of the key relevant comparisons are strict. Assumption 1(ii) characterizes a sufficient condition to ensure a unique stable matching.

Let $W^*(x)$ be the collection of stable payoff vectors $w^*(x)$ that support the unique stable matching $\mu^*(x)$.

As an example, consider the setting with two traits, $N = \{1, 2\}$. First, suppose $2f(1,2) < f(1,1) + f(2,2)$. The unique stable outcome involves positive assortative matching (PAM)—$\mu^*_{ii} = x_i/2$ for $i \in N$—and equal split—$w^*_i(x_i) = f(i,i)/2$ for $i \in N$. Next, suppose instead $2f(1,2) > f(1,1) + f(2,2)$. The stable outcome involves negative assortative matching (NAM). When $x_1 < 1/2$, the unique stable matching is given by $\mu^*_{12}(x) = x_1$ and $\mu^*_{22}(x) = x_2 - x_1$, and the unique stable payoff vector is given by $w^*(x) = (f(1,2) - f(2,2)/2, f(2,2)/2)$. When $x_1 > 1/2$, the unique stable matching is given by $\mu^*_{12}(x) = x_2$ and $\mu^*_{11}(x) = x_1 - x_2$, and the unique stable payoff vector is given by $w^*(x) = (f(1,1)/2, f(1,2) - f(1,1)/2)$. When $x_1 = 1/2$, the unique stable matching $\mu^*(x)$ is given by $\mu^*_{12}(x) = 1/2$, and any payoff vector $w^*(x)$ such that $f(i,i)/2 \leq w^*_i(x) \leq f(1,2) - f(j,j)/2$ for $i \in \{1,2\}$ and $j \neq i$, and $w^*_1(x) + w^*_2(x) = f(1,2)$ is stable.

2.2 The Evolution of Traits

The distribution of traits evolves over time, guided by the payoffs associated with different traits. We assume that the process of forming partnerships reaches a stable outcome much more rapidly than the evolution of traits. Thus, when the distribution of traits shifts, the stable outcome swiftly adapts.[2]

We consider the replicator dynamic to describe the evolution of traits. The replicator dynamic can be derived from a biological growth model (Taylor and Jonker, 1978), an intergenerational cultural transmission model (Bisin and Verdier, 2001), or learning models based on imitation (Sandholm, 2010). Let $\overline{w}^*(x) = \sum_{j \in N} x_j w^*_j(x)$ denote the average payoff in a stable outcome. Note

---

[2] This approach parallels the indirect evolutionary approach that studies preference evolution (Güth and Yaari, 1992; Güth, 1995), where agents' preferences evolve at a much slower pace compared to their behaviors.

that since there is a unit mass of agents, the average payoff equals the total payoff: $\overline{w}^*(x) = \sum_{i \in N} \sum_{j \leq i} \mu^*_{ij}(x) f(i,j)$. Since the unique stable matching $\mu^*(x)$ must maximize the total payoff, $\overline{w}^*(x)$ does not vary across $w^*(x) \in W^*(x)$. Let $\overline{\mathbf{w}}^*(x) = (\overline{w}^*(x), \overline{w}^*(x), ..., \overline{w}^*(x)) \in \mathbb{R}^n$ be the n-dimensional vector in which each of the n entries equals $\overline{w}^*(x)$. The replicator dynamic is given by a differential inclusion:[3]

$$\dot{x} \in V^{RDI}(x) \equiv \{x \cdot [w^*(x) - \overline{\mathbf{w}}^*(x)] \,|w^*(x) \in W^*(x)\}. \tag{RDI}$$

Hence, the fraction of agents who carry a trait that is associated with an above-average payoff grows and vice versa.

The dynamic described in (RDI) is a differential inclusion because of the possible indeterminacy of stable payoff vectors. When there are multiple stable payoff vectors, the path the dynamic takes depends on the realized stable payoff vector. For example, at x, if the realized stable payoff vector is $w^*(x)$, then the growth rate of trait i is given by $\dot{x}_i = x_i \cdot [w^*_i(x) - \overline{w}^*(x)]$ for any $i \in N$, which takes the standard form of the replicator dynamic.

We have the following theorem from Sandholm (2010) that provides sufficient conditions for the existence of solution trajectory.

Theorem 1 *(Sandholm, 2010, Theorem 6.A.1) Consider a differential inclusion $\dot{x} \in V(x)$. If V is nonempty, bounded, convex-valued, and upper-hemicontinuous, then for any initial condition $\xi \in \mathbb{R}^n$, there exists a* ***Carathéodory solution trajectory*** *$\{x^t\}_{t \in [0,T]}$ to the differential inclusion. That is, the solution trajectory is Lipschitz continuous, and its derivative $\dot{x}_t \in V(x)$ at almost all times $t \in [0,T]$.*

The following lemma shows that the replicator dynamic we consider is well-behaved and, thus, combined with Theorem 1, has a solution trajectory in the interior of X. Note that the uniqueness of the solution trajectory is not guaranteed.

---

[3] A differential inclusion is a generalization of a differential equation where the derivative of a function is constrained to belong to a set of possible values rather than being determined by a single value.

Lemma 2 *The correspondence V^{RDI} is nonempty-valued, bounded, convex-valued, and upper-hemicontinuous in X_{++}.*

2.3 Analysis

We now study the convergence and stability of the dynamic processes. Convergence informs us where the dynamic process ends up from all initial conditions, and stability determines whether the convergent states of the population are robust against perturbations. Please see Sandholm (2010) for a textbook treatment of these concepts and related results.

Definition 2 *The* ***ξ-limit set*** *$\omega(\xi)$ is the set of all population states such that the solution trajectory $\{x^t\}_{t\geq 0}$ starting from $x^0 = \xi$ approaches arbitrarily closely infinitely often. Formally,*

$$\omega(\xi) := \left\{ y \in X : \exists \{t_k\}_{k=1}^{\infty} \text{ with } \lim_{k\to\infty} t_k = \infty \text{ s.t. } x^0 = \xi \text{ and } \lim_{k\to\infty} x^{t_k} = y \right\}.$$

An element in a ξ-limit set is called a ***ξ-limit point****.*

The collection of all limit points is denoted by the set $\Omega = \cup_{\xi \in X} \omega(\xi)$: It captures the notion of recurrence of the dynamics. Let $A \subseteq X$ be a closed set, and call $O \subseteq X$ a neighborhood of A if it is open relative to X and contains A.

Definition 3 *Set $A \subset X$, a proper subset of X, is* ***Lyapunov stable*** *if for every neighborhood O of A, there exists a neighborhood O' of A such that every solution trajectory $\{x^t\}_{t\geq 0}$ that starts in O' is contained in O, that is, $x^0 \in O'$ implies that $x^t \in O$ for all $t \geq 0$. If A is not Lyapunov stable, we call it* ***Lyapunov unstable****.*

Intuitively, Lyapunov stability requires that any displacement from A does not lead the process to go "far" from A at any point in time.

Definition 4 *Set $A \subset X$ is* ***attracting*** *if there is a neighborhood $Y \subseteq X$ of A such that every solution trajectory that starts in Y converges to A, that is, $\xi \in Y$ implies $\omega(\xi) \subseteq A$. Set $A \subset X$ is*

globally attracting *if it is attracting with $Y = X$. The set of points $\xi \in X$ such that $\omega(\xi) \subseteq A$ is called the* ***basin of attraction*** *of A.*

Intuitively, that A is attracting requires that given any displacement from A, the process returns to A in the limit.

Definition 5 *Set $A \subset X$ is* ***asymptotically stable*** *if it is Lyapunov stable and attracting. Set $A \subset X$ is* ***globally asymptotically stable*** *if it is Lyapunov stable and globally attracting.*

Intuitively, asymptotic stability of A requires that, given any displacement from A, the process never travels "far" from A and returns to A in the limit. The limiting behavior of deterministic dynamics can be characterized as follows. We will focus on the case in which there exists a unique match that leads to the highest payoff.

Proposition 1 ***(i)*** *Suppose there exists a unique i such that $f(i,i) = \max_{k,l} f(k,l)$. The population state $x^* = e_i$, the ith standard basis vector, is the unique globally asymptotically stable state under the replicator dynamic with respect to $X \backslash \{x | x_i = 0\}$. It is also the efficient state.* ***(ii)*** *Suppose there exist i and $j \neq i$ such that $f(i,j) = \max_{k,l} f(k,l)$. The population state x^* such that $x_i^* = x_j^* = 0.5$ is the unique globally asymptotically stable state under the replicator dynamic with respect to $X \backslash \{x | x_i = 0$ or $x_j = 0$ or both$\}$. It is also the efficient state.*

Proposition 1 shows that the population states that maximize the average payoff in stable matching are asymptotically stable. See Appendix A for Theorems 2 and 3 that we use to prove stability and convergence based on the Lyapunov method. In the first case, when there exists a unique i such that $f(i,i) = \max_{k,l} f(k,l)$, we define $L^{RDI}(x) = -\log x_i$. In the second case, when there exist i and $j \neq i$ such that $f(i,j) = \max_{k,l} f(k,l)$, we define $L^{RDI}(x) = -[\log(2x_i) + \log(2x_j)]$. We will use $L^{RDI}(x)$ as the Lyapunov function for (RDI).

To make Proposition 1 more concrete, consider the setting with two productive traits. Without loss of generality, assume that $f(1,1) > f(2,2)$. Hence, the two traits differ in own-match productivity. When $2f(1,2) > f(1,1)+f(2,2)$, we call the payoff structure ***submodular***. There is always a unique stable matching in which the mass of cross-trait matching is maximized (negative assortative matching). When $2f(1,2) < f(1,1) + f(2,2)$, we call the payoff structure ***supermodular***. There is always a unique stable matching with only own-trait matching (PAM).

Case 1. Suppose $f(1,2) > f(1,1)$. This immediately implies submodularity. In this case, we always have a unique stable payoff as long as $x_1 \neq 1/2$. When $x_1 < 1/2$, $w_1(x) = f(1,2) - f(2,2)/2 > w_2(x) = f(2,2)/2$. When $x_1 > 1/2$, $w_1(x) = f(1,1)/2 < w_2(x) = f(1,2) - f(1,1)/2$. Hence, for (RDI), $\dot{x}_1 > 0$ when $0 < x_1 < 1/2$; $\dot{x}_1 < 0$ when $x_1 > 1/2$, which guarantees the globally asymptotic stability of $x^* = (1/2, 1/2)$ given any interior initial state.

Case 2. Suppose $f(1,1) > f(1,2)$ and $2f(1,2) > f(1,1) + f(2,2)$. We still have submodularity. When $x_1 < 1/2$, $w_1(x) = f(1,2) - f(2,2)/2 > w_2(x) = f(2,2)/2$. When $x_1 > 1/2$, $w_1(x) = f(1,1)/2 > w_2(x) = f(1,2) - f(1,1)/2$. Hence, for (RDI), $\dot{x}_1 > 0$ when $x_1 > 0$, which guarantees the globally asymptotic stability of $x_1 = 1$ in $(0,1]$.

Case 3. Suppose $f(1,1) > f(1,2)$ and $2f(1,2) < f(1,1) + f(2,2)$. In this case, we have supermodularity. We always have $w_1(x) = f(1,1)/2 > w_2(x) = f(2,2)/2$. Hence, for (RDI), $\dot{x}_1 > 0$ when $x_1 > 0$, which guarantees the globally asymptotic stability of $x_1 = 1$ in $(0,1]$.

To summarize this two-trait example, while supermodularity guarantees that trait 1, the more productive trait, prevails through evolution, submodularity does not guarantee the coexistence of both traits. Only when cross-trait match leads to the highest production can both traits coexist. These results confirm Proposition 1, the population state that maximizes the average payoff in stable matching is the long-run prediction of the evolutionary process. When $f(1,2)$ is the highest, $x = (1/2, 1/2)$ maximizes the average payoff. When $f(1,1)$ is the highest, $x = (1,0)$ maximizes the average payoff.

3 Sharing Norms and Efficiency

In this section, rather than letting agents determine surplus division endogenously within the market, we introduce exogenous sharing norms that specify how the surplus is distributed in each match. These norms reflect the cultural conventions that governed resource allocation in society before the emergence of free markets.

Let u_{ij} denote a trait-i agent's payoff from her match with a trait-j agent. $u_{ij} + u_{ji} = f(i, j)$. We adopt the stability concept in two-sided NTU (nontransferable utility) matching market (Gale and Shapley, 1962) to our setting of one-sided matching markets with a continuum of agents to characterize the stable matching of the partnership formation process.

Definition 6 *A* ***stable matching*** $\mu^*(x)$ *satisfies*

*(i) (**individual rationality**) each person gets at least as much as staying unmatched:* $u_{ij} \geq 0$ *for any* $i, j \in N$ *such that* $\mu^*_{ij}(x) > 0$*;*

*(ii) (**no blocking pair**) no pair of persons can both be strictly better off than what they have in their current matches:* $u_{ij} \leq u_{ik}$ *or* $u_{ji} \leq u_{jl}$ *for any* $i, j, k, l \in N$ *such that* $\mu^*_{ik}(x) > 0$ *and* $\mu^*_{jl}(x) > 0$.

Carmona and Laohakunakorn (2024) prove the existence of stable matching. For x and $\mu^*(x)$, let $u^*_i(x)$ denote the average payoff of all trait-i agents. Let $u^*(x) = (u^*_1(x), u^*_2(x), ..., u^*_n(x)) \in \mathbb{R}^n$ denote the vector of average payoffs. Let $\overline{u}^*(x)$ denote the average payoff in the entire population. Let $\overline{\mathbf{u}}^*(x) = (\overline{u}^*(x), \overline{u}^*(x), ..., \overline{u}^*(x)) \in \mathbb{R}^n$. The evolution of traits is still dictated by the replicator dynamic. But now it is given by a differential equation:

$$\dot{x} = V^{RDE}(x) \equiv \{x \cdot [u^*(x) - \overline{\mathbf{u}}^*(x)]\}\,. \tag{RDE}$$

Theorem 4 in Appendix A establishes asymptotic stability for (RDE).

3.1 Equal Sharing Norm

First, suppose two agents in a pair equally split the payoff they jointly produce. That is, $u_{ij} = f(i, j)/2$. Without loss of generality, assume $f(1, 1) > f(i, i)$ for any $i \neq 1$.

Proposition 2 ***(i)*** *Suppose $f(1,1) = \max_{k,l} f(k,l)$. Then $x^* = e_1$, where e_1 is the first standard basis vector, is globally asymptotically stable under* (RDE) *with respect to $X\backslash\{x|x_1 = 0\}$. It is also the efficient state.* ***(ii)*** *Suppose there exist i and j such that $i \neq j$ and $f(i,j) = \max_{k,l} f(k,l)$. Then x^* such that $x_i^* = x_j^* = 1/2$ is globally asymptotically stable under* (RDE) *with respect to $X\backslash\{x|x_i = 0$ or $x_j = 0$ or both$\}$. It is also the efficient state.*

Similar to the analysis for the market sharing norm, here we will focus on two cases. In the first case when a single trait (trait 1) is the most productive in terms of production, we define $L^{RDE}(x) = -\log x_1$. In the second case where there exist i and j such that $i \neq j$ and $f(i,j) = \max_{k,l} f(k,l)$, we define $L^{RDE}(x) = -[\log(2x_i) + \log(2x_j)]$. Proposition 2 will use $L^{RDE}(x)$ as the Lyapunov function for (RDE).

Note that $L^{RDE}(x)$ is identical to $L^{RDI}(x)$, which we use to prove stability for (RDI) in Section 2.3. Although they are the same, they function differently in the two settings due to the inherent differences between TU and NTU matching. Specifically, when a pair of traits (including potentially the same trait) are most productive when matched, these traits will pair with each other to the greatest extent possible under NTU matching, as in the equal sharing norm setting. This results in their average payoff being higher than the population average—an essential condition for the Lyapunov method to apply. However, under TU matching, as in the market sharing norm setting, these traits may not necessarily pair with each other to the greatest extent possible. Nevertheless, the no-blocking-pair condition under TU matching still ensures that the average payoff for these traits exceeds the population average.

In sum, under the equal sharing norm, efficiency is always achieved through evolution for any payoff structure f.

To illustrate, we consider settings with three traits. Suppose $f(1,1) > f(1,2)$, $f(1,1) > f(1,3)$, $f(2,2) > f(1,2)$, $f(2,2) < f(2,3)$, $f(3,3) < f(1,3)$, and $f(3,3) > f(2,3)$. This payoff structure leads to PAM and the globally attracting and stable state $x^* = (1,0,0)$. Figure 1a provides a graphic illustration of (RDE).

It is worth noting that, given the parameter values used to generate Figure 1a, the market sharing norm would not yield PAM. Instead, in the stable matching, all trait-1 agents are matched with each other, and traits 2 and 3 are cross-matched to the greatest extent possible. This matching generates a higher total payoff than PAM under the equal sharing norm, illustrating that equal sharing can be inefficient at the behavioral level. Nevertheless, under equal sharing the payoffs to traits 1, 2, and 3 are 3.25, 1, and 2, respectively. Under market sharing, the corresponding payoffs are 3.25, 1.5, and 2 when $x_2 < x_3$, and 3.25, 1, and 2.5 when $x_2 > x_3$. Hence, regardless of the sharing norm, the replicator dynamic converges to $x^* = (1, 0, 0)$, although the speed of convergence may differ across norms.

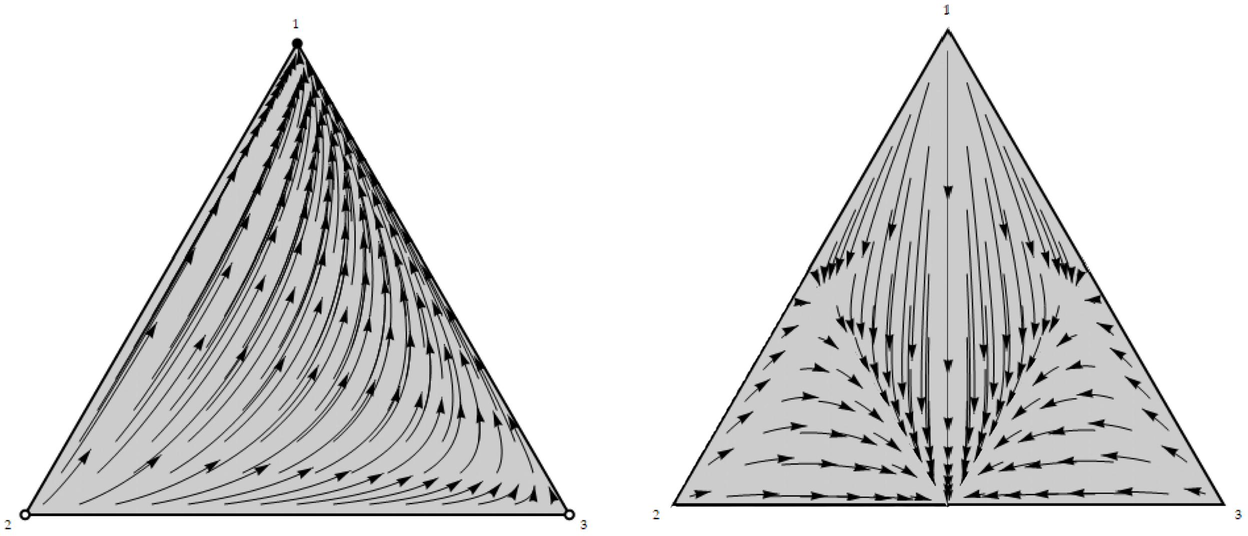
Alternatively, suppose $f(1,1) < f(1,2) < f(1,3)$, $f(2,2) < f(1,2) < f(2,3)$, and $f(3,3) < f(1,3) < f(2,3)$. Suppose the initial state $x^0 = (x_1^0, x_2^0, x_3^0) = (0.3, 0.5, 0.2)$. The unique stable matching is characterized by $\mu_{12} = 0.3, \mu_{23} = 0.2$. Then the payoffs under stable matching are $u_1^*(x^0) = f(1,2)/2$, $u_2^*(x^0) = 0.6 \cdot [f(1,2)/2] + 0.4 \cdot [f(2,3)/2]$, and $u_3^*(x^0) = f(2,3)/2$. The average payoff is $\overline{u}^*(x^0) = 0.6 \cdot [f(1,2)/2] + 0.4 \cdot [f(2,3)/2]$. Hence, under the replicator dynamic, $\dot{x}_1 = x_1^0 \cdot [u_1^*(x^0) - \overline{u}^*(x)] < 0$, $\dot{x}_2 = x_2^0 \cdot [u_2^*(x^0) - \overline{u}^*(x)] = 0$, and $\dot{x}_3 = x_3^0 \cdot [u_3^*(x^0) - \overline{u}^*(x)] > 0$. Hence, x_1 shrinks, x_2 stagnates, and x_3 grows. Therefore, $\lim_{t \to \infty} x^t = (0, 0.5, 0.5)$. Figure 1b provides a graphic illustration of (RDE).

3.2 Unequal Sharing Norms

Now suppose the society has a universal sharing norm that features an unequal division of joint production. Let $\alpha_{ij} \in (0,1)$ denote the share of joint production a trait-i agent receives when matched with a trait-j agent such that $\alpha_{ji} = 1 - \alpha_{ij}$.

Assume that the universally unequal sharing norm takes the following form: $\alpha_{ii} = 0.5$ for any i and $\alpha_{ij} > 0.5$ if $i < j$. The society is hierarchical in this case, where traits also represent hereditary social ranks and a higher ranked agent can receive a larger share when matched with a lower ranked agent. Assume $f(k,k) = \max_i f(i,i)$, which indicates that trait-k is the most productive in terms of matching with its own type. We allow $k \neq 1$ to maintain generality and avoid assuming that the highest ranked trait is also the most productive in own-trait matching.

Figure 1: Comparison of the replicator dynamic under PAM and Non-PAM given the equal sharing norm.



(a) Replicator dynamic under PAM.

(b) Replicator dynamic under Non-PAM.

Note: Panel (a) illustrates the replicator dynamic when $f(1,1) = 6.5$, $f(1,2) = 1$, $f(1,3) = 5$, $f(2,2) = 2$, $f(2,3) = 3.5$, and $f(3,3) = 4$. Panel (b) illustrates the replicator dynamics when $f(1,1) = 1$, $f(1,2) = 3$, $f(1,3) = 5$, $f(2,2) = 2$, $f(2,3) = 6$, and $f(3,3) = 4$.

For illustration, consider the following examples.

Suppose there are two traits: 1 and 2. The payoff structure satisfies that $f(1,1) = 6, f(2,2) = 8$ and $f(1,2) = 10$, and the sharing norm satisfies that $\alpha_{12} = 0.9$ and $\alpha_{21} = 0.1$. In this case, $u_{11} = 3, u_{22} = 4, u_{12} = 9, u_{21} = 1$, which induces PAM, so the evolutionary dynamic is simply driven by the comparison of u_{11} and u_{22}, which leads to a unique stable state $(0,1)$. However, this state is not efficient because society can achieve a higher average payoff in the state $(0.5, 0.5)$.

Suppose there are two traits 1 and 2. The payoff structure satisfies that $f(1,1) = 6, f(2,2) = 7$ and $f(1,2) = 10$, and the sharing norm satisfies that $\alpha_{12} = 0.6$ and $\alpha_{21} = 0.4$. In this case, $u_{11} = 3, u_{22} = 3.5, u_{12} = 6, u_{21} = 4$, which induces NAM. To determine the stable state(s), we compare the average payoffs of the trait-1 and trait-2 agents. When $x_1 \leq 0.5$,

$$u_1^*(x) = u_{12} = 6 > u_2^*(x) = \frac{x_1}{x_2}u_{21} + \frac{x_2 - x_1}{x_2}u_{22} = 4\frac{x_1}{x_2} + 3.5\frac{x_2 - x_1}{x_2}.$$

Hence, no state is stable in this region. When $x_1 > 0.5$,

$$u_1^*(x) = \frac{x_1 - x_2}{x_1}u_{11} + \frac{x_2}{x_1}u_{12} = 3\frac{x_1 - x_2}{x_1} + 6\frac{x_2}{x_1},\ u_2^*(x) = u_{21} = 4.$$

Hence, there exists a unique stable state $(3/4, 1/4)$. However, this state is not efficient, as society can achieve a higher average payoff in the state $(0.5, 0.5)$.

The examples demonstrate that universally unequal sharing norms can lead to inefficiency through evolution. We have the following general result.

Proposition 3 *For any universally unequal sharing norm, there exists a payoff structure f that leads to an asymptotically stable state that is inefficient.*

Next, we ask the question that for a given payoff structure f, whether there exists a universally unequal sharing norm that leads to an asymptotically stable state that is efficient. We obtain the following results.

Proposition 4 *(i) Suppose there exist i and $j \neq i$ such that $f(i,j) = \max_{k,l} f(k,l)$. Under any unequal sharing norm $\{\alpha_{ij}\}$, the efficient population state x^* with $x_i^* = x_j^* = 0.5$ is not stable.* ***(ii)*** *Suppose there exists a unique i such that $f(i,i) = \max_{k,l} f(k,l)$. There exists an unequal sharing norm such that the efficient population state $x^* = e_i$ is the unique globally asymptotically stable state and agents equally share their output.*

Proposition 4 shows that (i) when efficiency requires heterogeneity, unequal sharing norms necessarily lead to inefficiency; and (ii) when efficiency requires homogeneity, certain unequal sharing norms can still lead to efficiency, but in the efficient state, only one trait survives, and all agents sharing this trait equally share their outputs with their partners, so effectively no unequal sharing norm is observed. In other words, under a fixed unequal sharing norm, the evolutionarily stable outcome is either (i) an inefficient state with unequal sharing norm or (ii) an efficient state with a observationally equal sharing norm.

For illustration, consider the case with two productive traits.

Case 1. Suppose $f(1,2) > f(1,1)$ and $f(1,2) > f(2,2)$. If $f(1,1)/2 > f(1,2)\alpha_{12}$ or $f(2,2)/2 > f(1,2)\alpha_{21}$ or both, then we have PAM and (RDE) converges to a homogeneous state which must be inefficient. If $f(1,2)\alpha_{12} > f(1,1)/2$ and $f(1,2)\alpha_{21} > f(2,2)/2$, then we have NAM. However, at the efficient state x^* such that $x_1^* = x_2^* = 1/2$, $u_1(x^*) = f(1,2)\alpha_{12} \neq f(1,2)\alpha_{21} = u_2(x^*)$, implying x^* is unstable under (RDE).

Case 2. Suppose $f(1,1) > f(1,2)$ or $f(2,2) > f(1,2)$ or both. Also assume that $f(1,1) \neq f(2,2)$. Let $\alpha_{12} = 1/2 + \varepsilon$ and $\alpha_{21} = 1/2 - \varepsilon$, for some $\varepsilon \in (0, 1/2)$. Then there always exists ε such that $f(1,1)/2 > f(1,2)(1/2 + \varepsilon)$ or $f(2,2)/2 > f(1,2)(1/2 - \varepsilon)$ or both. Such ε induces PAM, and leading (RDE) to the efficient state e_1 (resp. e_2) if $f(1,1) > f(2,2)$ (resp. $f(2,2) > f(1,1)$). This demonstrates that as long as ε is sufficiently small (i.e., the unequal sharing norm is close to equal sharing), efficiency can be sustained in this case.

4 Conclusions

This paper offers an evolutionary perspective on the resilience of equal sharing norms and challenges the assumption that market-based approaches are inherently superior. Through a two-tier evolutionary framework, we demonstrate how equal sharing, despite its rigidity, supports the evolution of productive traits that facilitate efficient partnerships and thereby enhance long-term productivity. Our findings suggest that the equal sharing norm, far from being an outdated relic, continues to thrive alongside market mechanisms due to its capacity to shape stable and productive populations. We also show that a slight deviation from equal sharing, namely an unequal sharing norm close to equal sharing, may persist but only when efficiency requires the population to be homogeneous in productive traits.

The current model focuses solely on pairwise matching, resulting in a stable distribution of preferences that is either monomorphic or bimorphic. A natural extension would be to consider team formation. We anticipate that our results will generalize naturally to any fixed team size and stable distributions can exhibit richer trait diversity. It would also be an intriguing direction to explore more general coalition formation settings.

A Omitted Proofs

Proof of Lemma 1. The set of feasible one-sided matchings is

$$\mathcal{M}_1(x) = \left\{ \mu = (\mu_{ij}) \,\middle|\, 2\mu_{ii} + \sum_{j \neq i} \mu_{ij} \leq x_i \,\forall i \in N \text{ and } \mu_{ij} = \mu_{ji} \,\forall i, j \in N \right\}.$$

We define an outcome (μ, w) as the specification of a feasible roommate matching μ and an associated vector of payoffs $(w_i)_{i \in N}$, with w_i denoting the payoff of each individual of trait i. These payoffs have to be feasible: The sum of payoffs across the population has to be equal to the total output under the matching μ. In a one-sided matching μ, the total surplus created is

$$\overline{w}^*(\mu) = \sum_{i \in N} \mu_{ii} f(i,i) + \sum_{i \in N} \left[\sum_{j \in N \setminus \{i\}} \frac{\mu_{ij}}{2} f(i,j) \right] = \sum_{i \in N} \left[\mu_{ii} f(i,i) + \sum_{j \in N \setminus \{i\}} \frac{\mu_{ij}}{2} f(i,j) \right].$$

The maximum of the aggregate surplus over the set of feasible one-sided matchings $\mathcal{M}_1(x)$ is

$$P_1(x) = \max_{\mu \in \mathcal{M}_1(x)} \overline{w}^*(\mu). \tag{P_1}$$

Explicitly, this ***one-sided primal program*** is

$$P_1(x) = \max_{\mu} \; \overline{w}^*(\mu)$$

$$\text{subject to} \qquad 2\mu_{ii} + \sum_{j \in N \setminus \{i\}} \mu_{ij} \leq x_i \quad \forall i \in N, \tag{feasibility}$$

$$\mu_{ij} = \mu_{ji} \quad \forall i, j \in N. \tag{symmetry}$$

Define the following ***one-sided dual program***:

$$D_2(x) = \min_{w = (w_i)_{i \in N}, A = (A_{ij})_{i,j \in N}} \sum_{i \in N} x_i w_i \tag{D_1}$$

$$\text{subject to} \qquad w_i \geq 0 \quad \forall i \in N \tag{individual rationality}$$

$$w_i + w_j \geq f(i,j) + A_{ij} \quad \forall i,j \in N, \tag{no blocking pair}$$

$$A_{ij} = -A_{ji} \quad \forall i,j \in N. \tag{anti-symmetry}$$

Consider an associated two-sided matching problem, which will be important for our subsequent proofs to leverage the well-established tools in the two-sided matching literature. Denote the set of feasible two-sided matchings by

$$\mathcal{M}_2(x,x) = \left\{ \nu = (\nu_{ij}) \middle| \sum_{j \in N} \nu_{ij} \leq x_i \, \forall i \in N \text{ and } \sum_{i \in N} \nu_{ij} \leq x_j \, \forall j \in N \right\},$$

and the total surplus of a two-sided matching ν by

$$\overline{w}_2^*(\nu) = \sum_{i,j \in N} \nu_{ij} f(i,j).$$

A stable outcome in the two-sided matching setting is defined analogously: (ν, u, v) is a ***two-sided stable outcome*** that consists of a feasible two-sided matching $\nu = (\nu_{ij})_{i,j \in N}$ and two-sided payoff vectors $u = (u_i)_{i \in N}$ and $v = (v_j)_{j \in N}$ if (i) (individual rationality) $u_i \geq 0$ for any $i \in N$ and $v_j \geq 0$ for any $j \in N$, (ii) (no blocking pair) $u_i + v_j \geq f(i,j)$ for any $i, j \in N$, and (iii) (surplus efficiency) $u_i + v_j = f(i,j)$ if $\nu_{ij} > 0$.

The two-sided surplus maximization problem of population state $(x/2, x/2)$—the ***two-sided primal program***—associated with the one-sided primal program (P_1) of population state x is

$$P_2(x/2, x/2) = \max_{\nu \in \mathcal{M}_2(x/2, x/2)} \overline{w}_2^*(\nu). \tag{P_2}$$

Explicitly,

$$P_2(x/2, x/2) = \max_{\nu} \overline{w}_2^*(\nu)$$

$$\text{subject to} \qquad \sum_{j \in N} \nu_{ij} \leq x_i \forall i \in N \text{ and } \sum_{i \in N} \nu_{ij} \leq x_j \forall j \in N. \tag{feasibility}$$

Note that the symmetry constraints in the one-sided primal program are absent in the two-sided primal program. The ***two-sided dual program*** is

$$D_2(x/2, x/2) = \min_{u,v} \sum_{i \in N} \frac{1}{2} x_i u_i + \sum_{j \in N} \frac{1}{2} x_j v_j \tag{D_2}$$

subject to $u_i \geq 0 \; \forall i \in N$ and $v_j \geq 0 \; \forall j \in N,$ (individual rationality)

$u_i + v_j \geq f(i, j) \quad \forall i, j \in N.$ (no blocking pair)

By Shapley and Shubik (1972) and Gretsky, Ostroy, and Zame (1992), there exist two-sided stable matchings ν and they coincide with the solution of the two-sided primal program (P_2). Moreover, the associated payoffs (u, v) solve the two-sided dual program (D_2). Finally, for any stable matching ν, $\nu_{ij} > 0$ implies $u_i + v_j = f(i, j)$, $\nu_{i0} > 0$ implies $u_i = 0$, and ν_{0j} implies $v_j = 0$.

Note that the two-sided matching problem is a particular case of the one-sided matching problem: If $f(i, j) = -\infty$ whenever i and j belong to the same side, then any optimal or stable matching will be bipartite.

Claim: The values of the one-sided primal program with population state $x \in X_{++}$ and the associated two-sided primal program with population state $(x/2, x/2)$ coincide:

$$P_1(x) = P_2(x/2, x/2).$$

Consider a matching $\mu \in \mathcal{M}_1(x)$. Define ν such that $\nu_{ij} = \mu_{ij}/2$ if $i \neq j$ and $\nu_{ii} = \mu_{ii}$. For any $i \in N$,

$$\sum_{j \in N} \nu_{ij} = \nu_{ii} + \sum_{j \neq i} \nu_{ij} = \mu_{ii} + \sum_{j \neq i} \mu_{ij}/2 \leq x_i/2,$$

and

$$\sum_{j \in N} \nu_{ji} = \nu_{ii} + \sum_{j \neq i} \nu_{ji} = \mu_{ii} + \sum_{j \neq i} \mu_{ji}/2 \leq x_i/2,$$

so $\nu \in \mathcal{M}_2(x/2, x/2)$, as the feasibility constraint for $\nu \in \mathcal{M}_2(x/2, x/2)$ is satisfied.

$$\overline{w}^*(\mu) = \sum_{i \in N} \mu_{ii} f(i,i) + \sum_{i \in N} \left[\sum_{j \in N \setminus \{i\}} \frac{\mu_{ij}}{2} f(i,j) \right] = \sum_{i \in N} \left[\sum_{j \in N} \nu_{ij} f(i,j) \right] = \overline{w}_2^*(\nu).$$

Hence, ν, a feasible two-sided matching in $(x/2, x/2)$, achieves the maximum surplus of one-sided matching under population state x (achieved by the optimal solution μ). We have

$$P_1(x) \le P_2(x/2, x/2). \tag{1}$$

Conversely, let (ν_{ij}) maximize aggregate surplus over $\mathcal{M}_2(x/2, x/2)$. By the symmetry of f, (ν_{ji}) is also a maximizer. Since (P_2) is a linear program, $\nu'_{ij} = (\nu_{ij} + \nu_{ji})/2$ also maximizes it. Define μ such that $\mu'_{ij} = 2\nu'_{ij}$ if $i \neq j$ and $\mu'_{ii} = \nu'_{ii}$. Then for any $i \in N$,

$$2\mu'_{ii} + \sum_{j \neq i} \mu'_{ij} = 2\nu_{ii} + \sum_{j \neq i} (\nu_{ij} + \nu_{ji}) = \sum_{j \in N} \nu_{ij} + \sum_{j \in N} \nu_{ji} \le x_i,$$

where $\sum_{j \in N} \nu_{ij} \le x_i/2$ and $\sum_{j \in N} \nu_{ji} \le x_i/2$ follow from the feasibility constraints of $\nu \in \mathcal{M}_2(x/2, x/2)$. Hence, $\mu' \in \mathcal{M}_1(x)$. The aggregate surplus under one-sided matching μ' is

$$\sum_{i \in N} \mu'_{ii} f(i,i) + \sum_{i \in N} \left[\sum_{j \neq i} \frac{\mu'_{ij}}{2} f(i,j) \right] = \sum_{i \in N} \nu'_{ii} f(i,i) + \sum_{i \in N} \left[\sum_{j \neq i} \nu'_{ij} f(i,j) \right] = \sum_{i \in N} \sum_{j \in N} \nu'_{ij} f(i,j).$$

Hence, μ', a feasible one-sided matching in population state x, achieves the maximal two-sided surplus in population state $(x/2, x/2)$ (achieved by the optimal solution ν'). Hence,

$$P_1(x) \ge P_2(x/2, x/2). \tag{2}$$

Combining (1) and (2), we have $P_1(x) = P_2(x/2, x/2)$.

By the claim, the values of one-sided primal program (P_1) $P_1(x)$ and two-sided primal program (P_2) $P_2(x/2, x/2)$ coincide. Let $\nu = (\nu_{ij})$ denote a solution of the two-sided primal program. By the duality result of Shapley and Shubik (1972) and Gretsky, Ostroy, and Zame (1992), there exist two-sided stable payoff vectors (u_i) and (v_j) that solve the two-sided dual program (D_2) $D_2(x/2, x/2)$, namely, such that $u_i \geq 0$ for any $i \in N$, $v_j \geq 0$ for any $j \in N$, $u_i + v_j \geq f(i,j)$ for any $i, j \in N$, and $u_i + v_j = f(i,j)$ whenever $\nu_{ij} > 0$.

Define $\mu_{ij} = \nu_{ij} + \nu_{ji}$ when $i \neq j$, $\mu_{ii} = \nu_{ii}$, and $w_i = (u_i + v_i)/2$ for any $i \in N$. We have $w_i \geq 0$, $w_i + w_j \geq f(i,j)$, $w_i + w_j = f(i,j)$ whenever $\mu_{ij} > 0$, and

$$\sum_{i \in N} x_i w_i = \frac{1}{2}\left[\sum_{i \in N} x_i u_i + \sum_{i \in N} x_i v_i\right] = P_2(x/2, x/2) = P_1(x) = \overline{w}^*(\mu).$$

Hence, (μ, w) is a one-sided stable outcome. ■

Proof of Lemma 2. First, fix a population state $x \in X_{++}$. The set of stable one-sided payoffs, $W^*(x)$, is convex. Fix two stable payoff vectors $w^\spadesuit \in W^*(x)$ and $w^\diamondsuit \in W^*(x)$. Take any $\lambda \in (0,1)$. We have (i) $\lambda w_i^\spadesuit + (1-\lambda) w_i^\diamondsuit \geq 0$ for any $i \in N$, because $w_i^\spadesuit \geq 0$ and $w_i^\diamondsuit \geq 0$ for any $i \in N$, (ii) $[\lambda w_i^\spadesuit + (1-\lambda) w_i^\diamondsuit] + [\lambda w_j^\spadesuit + (1-\lambda) w_j^\diamondsuit] \geq f(i,j)$ for any $i, j \in N$, because $w_i^\spadesuit + w_j^\spadesuit \geq f(i,j)$ and $w_i^\diamondsuit + w_j^\diamondsuit \geq f(i,j) \geq f(i,j)$.

Second, fix a population state $x \in X_{++}$. The set of stable one-sided payoffs, $W^*(x)$, is compact. Since $W^*(x)$ is a set in $\mathbb{R}^{|I|}$, it suffices to show that $W^*(x)$ is closed and bounded. Each element of $W^*(x)$ is bounded due to the surplus function being bounded. The set is also closed, because all individual rationality and no-blocking-pair constraints are non-strict inequalities.

Third, the correspondence, $W : X_{++} \rightrightarrows \mathbb{R}^{|I|}$, that specifies the set of stable one-sided payoffs $W^*(x)$ for population state $x \in X_{++}$, is upper-hemicontinuous. Proof: $W : X_{++} \rightrightarrows \mathbb{R}^{|I|}$ is upper-hemicontinuous if and only if $x^k \to x$, $w^k \in W(x^k)$, and $w^k \to w$ imply $w \in W^*(x)$. It suffices to show that $\overline{w}^*(x) = \sum_{i \in N} w_i x_i$ and that w satisfies $w_i \geq 0$ for any $i \in N$ and $w_i + w_j \geq f(i,j)$ for any $i, j \in N$. First, we show that $\overline{w}^*(x) = \sum_{i \in N} w_i x_i$.

$$\left|\sum_{i \in N} w_i x_i - \overline{w}^*(x)\right|$$

$$
\begin{aligned}
&= \left| \sum_{i\in N} w_i x_i - \sum_{i\in N} w_i^k x_i^k + \overline{w}^*(x^k) - \overline{w}^*(x) \right| \\
&\leq \left| \sum_{i\in N} w_i x_i - \sum_{i\in N} w_i^k x_i^k \right| + \left| \overline{w}^*(x^k) - \overline{w}^*(x) \right| \\
&\leq \left| \sum_{i\in N} w_i x_i - \sum_{i\in N} w_i x_i^k \right| + \left| \sum_{i\in N} w_i x_i^k - \sum_{i\in N} w_i^k x_i^k \right| + \left| \overline{w}^*(x^k) - \overline{w}^*(x) \right| \\
&\leq \left| \sum_{i\in N} w_i x_i - \sum_{i\in N} w_i x_i^k \right| + \left|\left| w - w^k \right|\right| + \left| \overline{w}^*(x^k) - \overline{w}^*(x) \right|.
\end{aligned}
$$

Fix ϵ. Since $x^k \to x$, there is a $k_1(\epsilon)$ such that all $k > k_1(\epsilon)$, $\left|\sum_{i\in N} w_i x_i - \sum_{i\in N} w_i x_i^k\right| \leq \epsilon/3$. Since $w^k \to w$, there is a $k_2(\epsilon)$ such that for all $k > k_2(\epsilon)$, $\left|\left|w - w^k\right|\right| \leq \epsilon/3$. Since $x^k \to x$, $\left|\overline{w}^*(x^k) - \overline{w}^*(x)\right| \leq \epsilon/3$. Therefore, $\left|\sum_{i\in N} w_i x_i - \overline{w}^*(x)\right| \leq \epsilon$ for any $\epsilon > 0$.

We also need to show that w satisfies the stability conditions. Suppose otherwise: $w_i + w_j < f(i,j)$ for a positive measure of (i,j) or $w_i < 0$ for some $i \in N$. Then $\lim_{k\to\infty} \|w - w^k\| > 0$, a contradiction to the assumption that $w^k \to w$. ■

In what follows, we provide the theorems we will depend on to use the Lyapunov method to show convergence and stability of the dynamics.

Theorem 2 *(Sandholm, 2010, Theorem 7.B.2) Let $A \subseteq X$ be closed and let $Y \subseteq X$ be a neighborhood of A. Let $L : Y \to R+$ be Lipschitz continuous with $L^{-1}(0) = A$. If each solution $\{x^t\}_{t\geq 0}$ of $\dot{x} \in V^{RDI}(x)$ (or $\dot{x} = V^{RDE}(x)$) satisfies $\dot{L}(x^t) \leq 0$ for almost all $t \geq 0$, then A is Lyapunov stable.*

Theorem 2 states that the existence of a Lipschitz continuous Lyapunov function $L(x)$ guarantees that $L^{-1}(0)$, the set of population states that minimize $L(x)$, is Lyapunov stable.

Theorem 3 *(Sandholm, 2010, Theorem 7.B.4) Let $Y \subset X$ be relatively open and inescapable under (RDI). Let $L : Y \to R$ be C^1 and satisfy (i) $\frac{\partial L}{\partial v}(x) \equiv \nabla L(x)'v \leq 0$ for all $v \in V^{RDI}(x)$ and $x \in Y$ and (ii) $\left[\mathbf{0} \notin V^{RDI}(x) \implies \frac{\partial L}{\partial v}(x) < 0\right]$ for all $v \in V^{RDI}(x)$ and $x \in Y$. Then, for all solutions $\{x^t\}$ of (RDI) with $x_0 \in Y$, $\omega(\{x^t\}) \subseteq \{x \in Y : \mathbf{0} \in V(x)\}$.*

Theorem 3 shows that the existence of a C^1 Lyapunov function $L(x)$ guarantees the solution trajectories converge to the states that minimize $L(x)$. Combining Theorems 2 and 3 gives us globally asymptotic stability of the set of population states that minimize $L(x)$ under (RDI).

Proof of Proposition 1. In the first case, $L^{RDI}(x) = -\log x_i$ is a C^1 function. Moreover, for any x such that $x_i > 0$, and for any $v = x \cdot [w^*(x) - \overline{w}^*(x)]$, we have

$$\begin{aligned} \frac{\partial L^{RDI}}{\partial v}(x) &= \nabla L^{RDI}(x)'v \\ &= -\left(\frac{1}{x_i}\right) x_i[w_i^*(x) - \overline{w}^*(x)] \\ &= -[w_i^*(x) - \overline{w}^*(x)]. \end{aligned}$$

By the definition of stable matching, $w_i^*(x) \geq f(i,i)/2$. By the fact that $f(i,i) = \max_{k,l} f(k,l)$, $f(i,i)/2 > \overline{w}^*(x)$ as long as $x \neq e_i$. Therefore, we have $\frac{\partial L^{RDI}}{\partial v}(x) < 0$. At $x^* = e_i$, the stable matching is given by $\mu_{ii}(e_i) = 0.5$, where $w_i^*(e_i) = \overline{w}^*(e_i) = f(1,1)/2$. Hence, $V^{RDI}(e^1) = \{\mathbf{0}\}$. Also, $L^{-1}(0) = \{e_i\}$. By Theorems 2 and 3, e_i is the globally asymptotically stable state with respect to $X \backslash \{x | x_i = 0\}$. In addition, $f(i,i) = \max_{k,l} f(k,l)$ implies that e_i is the unique efficient state.

In the second case, $L^{RDI}(x) = -[\log(2x_i) + \log(2x_j)]$ is a C^1 function. Moreover, for any x such that $x_i > 0$ and $x_j > 0$, and for any $v = x \cdot [w^*(x) - \overline{w}^*(x)]$, we have

$$\begin{aligned} \frac{\partial L^{RDI}}{\partial v}(x) &= \nabla L^{RDI}(x)'v \\ &= -\left(\frac{1}{x_i}\right) x_i[w_i^*(x) - \overline{w}^*(x)] - \left(\frac{1}{x_j}\right) x_j[w_j^*(x) - \overline{w}^*(x)] \\ &= -[(w_i^*(x) + w_j^*(x)) - 2\overline{w}^*(x)]. \end{aligned}$$

By the definition of stable matching, $w_i^*(x) + w_j^*(x) \geq f(i,j)$. By the fact that $f(i,j) = \max_{k,l} f(k,l)$, we have $f(i,j)/2 > \overline{w}^*(x)$. Therefore, we have $\frac{\partial L^{RDI}}{\partial v}(x) < 0$. At x^* such that $x_i^* = x_j^* = 0.5$, the stable matching is given by $\mu_{ij}(x^*) = 0.5$, where $w^*(x^*)$ that satisfies $w_i^*(x^*) = w_j^*(x^*) = f(i,j)/2$ is a possible stable payoff vector. Hence $\mathbf{0} \in V^{RDI}(x^*)$. Also,

$L^{-1}(0) = \{x^*\}$. By Theorems 2 and 3, x^* is the globally asymptotically stable state with respect to $X \backslash \{x | x_i = 0 \text{ or } x_j = 0 \text{ or both}\}$. In addition, $f(i,j) = \max_{k,l} f(k,l)$ implies that x^* is the unique efficient state. ■

Theorem 4 *(Sandholm, 2010, Corollary 7.B.6) Let $A \subseteq X$ be closed and let $Y \subseteq X$ be a neighborhood of A. Let $L : Y \to R$ be C^1 with $L^{-1}(0) = A$. If $\dot{L}(x) < 0$ for all $x \in Y - A$, then A is asymptotically stable under $\dot{x} = V^{RDE}(x)$. If in addition, $Y = X$, then A is globally asymptotically stable under $\dot{x} = V^{RDE}(x)$.*

Theorem 4 gives us globally asymptotic stability of the set of population states that minimize $L(x)$ under (RDE).

Proof of Proposition 2. In the first case, $f(1,1) = \max_{k,l} f(k,l)$ implies that $f(1,1) > f(1,i)$ for any $i \neq 1$. Hence, as long as $x_1 > 0$, all trait-1 agents must match among themselves and $u_1^*(x) = u_{11} = f(1,1)/2$ is the highest payoff an agent can obtain in the population.

We want to show that $L^{RDE}(x) = -\log x_1$ serves as a Lyapunov function for the replicator dynamic. First, it is a C^1 function with $L^{RDE}(x^*) = 0$. Furthermore, for any x^t such that $x_1^t > 0$,

$$\begin{aligned} dL^{RDE}(x^t)/dt &= \triangledown L^{RDE}(x^t)'\dot{x}^t \\ &= -\left(\frac{1}{x_1^t}\right) x_1^t [u_1^*(x^t) - \overline{u}^*(x^t)] \\ &= \overline{u}^*(x^t) - u_1^*(x^t) \\ &< 0 \text{ as long as } x^t \neq e_1. \end{aligned}$$

Hence, according to Theorem 4, x^* is globally asymptotically stable with respect to $X \backslash \{x | x_1 = 0\}$.

In the second case, $f(i,j) = \max_{k,l} f(k,l)$ implies that $f(i,j) > f(i,i)$ and $f(i,j) > f(j,j)$. Hence, these two traits of agents will match with each other to the greatest extent possible. Without loss of generality, assume that $0 < x_i^0 < x_j^0$. Then all trait-i agents are matched with trait-j agents, while the remaining trait-j agents match otherwise. In this case, $u_i^*(x^0) = u_{ij} = f(i,j)/2$, which is the largest payoff an agent can get.

Let $r(t) = x_i^t/x_j^t$. Observe that under the replicator dynamic, $\frac{dr(t)}{dt} = \frac{x_i^t}{x_j^t}[u_i^*(x^t) - u_j^*(x^t)] > 0$ as long as $x_i^t < x_j^t$ so that $u_i^*(x^t) - u_j^*(x^t) > 0$. Also, once $x_i^t = x_j^t$, $r(t) = 1$ and $u_i^*(x^t) = u_j^*(x^t) = f(i,j)/2$ because all trait-i agents are exactly matched with all trait-j agents. Hence, $\frac{dr(t)}{dt} = 0$. In sum, $r(t)$ is a strictly increasing function bounded by 1, implying that $\lim_{t\to\infty} r(t) = 1$. Define $Y = \{x|x_i = x_j > 0\}$. This part of the proof shows that the dynamic always converges to Y from $X\backslash\{x|x_i = 0 \text{ or } x_j = 0 \text{ or both}\}$.

Next, we want to show that x^* is asymptotically stable with respect to Y. $L^{RDE}(x) = -[\log(2x_j) + \log(2x_i)]$ is a C_1 function in Y with $L^{RDE}(x^*) = 0$. In addition,

$$
\begin{aligned}
dL^{RDE}(x^t)/dt &= \bigtriangledown L^{RDE}(x^t)'\dot{x}^t \\
&= -\left(\frac{1}{2x_i^t}\right) x_i^t \left[u_i^*(x^t) - \overline{u}^*(x^t)\right] - \left(\frac{1}{2x_j^t}\right) x_j^t \left[u_j^*(x^t) - \overline{u}^*(x^t)\right] \\
&= \overline{u}^*(x^t) - \left[u_i^*(x^t) + u_j^*(x^t)\right]/2 \\
&= \overline{u}^*(x^t) - f(i,j)/2 \\
&< 0 \text{ as long as } x^t \neq x^*.
\end{aligned}
$$

Hence, $L^{RDE}(x)$ is a Lyapunov function for the replicator dynamic on Y and once the dynamic reaches Y, the dynamic converges to x^* according to Theorem 4. In sum, we can conclude that x^* is globally asymptotically stable with respect to $X\backslash\{x|x_i = 0 \text{ or } x_j = 0 \text{ or both}\}$. ■

Proof of Proposition 3. We will prove by construction. First, let $f(i,i) = c$ for some constant $c > 0$, for any $i \neq 1$. For any $1 < i < j$, let $f(i,j) = c + \nu_{ij}$, where $0 < \nu_{ij} < c[1/(2\alpha_{ji}) - 1]$. We must have that for any $1 < i < j$,

$$u_{ji} = \alpha_{ji}(c + \nu_{ij}) < u_{jj} = u_{ii} = c/2 < u_{ij} = \alpha_{ij}(c + \nu_{ij}).$$

This implies that in any stable matching, no $i - j$ match will be formed.

Next, let $f(1,i) = c + \epsilon_i$ for any $i \neq 1$, where $0 < \epsilon_i < c[1/(2\alpha_{i1}) - 1]$. Let $\epsilon = \min_i \epsilon_i$ and let $f(1,1) = c + \epsilon/2$. Then we have, for any $i \neq 1$,

$$u_{i1} = \alpha_{i1}(c + \epsilon_i) < c/2 = u_{ii} < u_{11} = (c + \epsilon/2)/2 < (c + \epsilon_i)/2 < \alpha_{1i}(c + \epsilon_i) = u_{1i}.$$

This implies that in any stable matching, no $1 - i$ match will be formed. Hence, the unique stable matching must be PAM and e_1 is the globally asymptotically stable state $X \backslash \{x | x_1 = 0\}$. However, it is not efficient because $f(1,1) < f(1,i)$ for any $i \neq 1$. ■

Proof of Proposition 4. In the first case, $f(i,j) = \max_{k,l} f(k,l)$ implies that the efficient state must take the form of x^* with $x_i^* = x_j^* = 0.5$. However, since $\alpha_{ij} \neq \alpha_{ji}$, $u_i(x^*) = f(i,j)\alpha_{ij} \neq f(i,j)\alpha_{ji} = u_j(x^*)$. Hence, x^* is unstable.

In the second case, $f(i,i) = \max_{k,l} f(k,l)$ implies that $f(i,i) > f(k,l)$ for any k, l such that not both of them equal i. If $k > l$, since $\alpha_{kl} < 1/2$, we must have $f(i,i)/2 > f(k,l)\alpha_{kl}$. If $k = l$, then $f(i,i)/2 > f(k,k)/2$. If $k < l$, since $f(i,i) > f(k,l)$, there exists $\alpha_{kl} > 1/2$ such that $f(i,i)/2 > f(k,l)\alpha_{kl}$. Hence, given such α_{kl}, as long as $x_i > 0$, all trait-i agents must match among themselves and $u_i^*(x) = u_{ii} = f(i,i)/2$ is the highest payoff an agent can obtain in the population. Then we can repeat the same argument as in Proposition 2 to show that $x^* = e_i$ is globally asymptotically with respect to $X \backslash \{x | x_i = 0\}$. ■

References

Abramitzky, Ran. 2008. “The Limits of Equality: Insights from the Israeli Kibbutz.” *Quarterly Journal of Economics* 123 (3):1111–1159.

———. 2011. “Lessons from the Kibbutz on the Equality-Incentives Trade-Off.” *Journal of Economic Perspectives* 25 (1):185–208.

———. 2018. *The Mystery of the Kibbutz: Egalitarian Principles in a Capitalist World*. Princeton, New Jersey: Princeton University Press.

Azevedo, Eduardo M., E. Glen Weyl, and Alexander White. 2013. “Walrasian Equilibrium in Large, Quasilinear Markets.” *Theoretical Economics* 8 (2):281–290.

Becker, Gary S. 1973. “A Theory of Marriage: Part I.” *Journal of Political Economy* 81 (4):813–846.

Bisin, Alberto and Thierry Verdier. 2001. “The Economics of Cultural Transmission and the Dynamics of Preferences.” *Journal of Economic Theory* 97 (2):298–319.

Carmona, Guilherme and Krittanai Laohakunakorn. 2024. “Stable Matching in Large Markets with Occupational Choice.” *Theoretical Economics* 19:1261–1304.

Chiappori, Pierre-André, Bernard Salanié, and Alfred Galichon. 2019. “On Human Capital and Team Stability.” *Journal of Human Capital* 13 (2):236–259.

Echenique, Federico, Sangmok Lee, Matthew Shum, and Bumin M. Yenmez. 2013. “The Revealed Preference Theory of Stable and Extremal Stable Mathcings.” *Econometrica* 81 (1):153–171.

Gale, David and Lloyd S. Shapley. 1962. “College Admissions and the Stability of Marriage.” *The American Mathematical Monthly* 69 (1):9–15.

Gretsky, Neil E., Joseph M. Ostroy, and William R. Zame. 1992. “The Nonatomic Assignment Model.” *Economic Theory* 2:103–127.

Güth, Werner. 1995. “An Evolutionary Approach to Explaining Cooperative Behavior by Reciprocal Incentives.” *International Journal of Game Theory* 24:323–344.

Güth, Werner and Menahem Yaari. 1992. “An Evolutionary Approach to Explain Reciprocal Behavior in a Simple Strategic Game.” In *Explaining Process and Change - Approaches to Evolutionary Economics*, edited by Ulrich Witt. University of Michigan Press, 23–34.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath. 2001. “In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies.” *American Economic Review* 91 (2):73–78.

Hiller, Victor, Jiabin Wu, and Hanzhe Zhang. 2023. “Marital Preferences and Stable Matching in Cultural Evolution.” *Journal of Economic Theory* 210:105671.

Lee, Richard B. 1979. *The !Kung San: Men, Women, and Work in a Foraging Society*. Cambridge: Cambridge University Press.

Polcz, Sarah. 2023. “Loyalties v. Royalties.” *Hastings Law Journal* 74 (3):765–822.

Sandholm, William H. 2010. *Population Games and Evolutionary Dynamics*. The MIT Press.

Shapley, Lloyd S. and Martin Shubik. 1972. “The Assignment Game I: The Core.” *International Journal of Game Theory* 1 (1):111–130.

Skyrms, Brian. 2014. *Evolution of the Social Contract*. Cambridge University Press.

Smith, Thomas A. 2001. “Equality, Evolution and Partnership Law.” *Journal of Bioeconomics* 3:99–121.

Taylor, Peter D. and Leo B. Jonker. 1978. “Evolutionarily Stable Strategies and Game Dynamics.” *Mathematical Biosciences* 40:145–156.

Wang, Ziwei and Jiabin Wu. 2025. “Preference Evolution under Partner Choice.” Working Paper.

Woodburn, James. 1982. “Egalitarian Societies.” *Man* 17 (3):431–451.

Wu, Jiabin and Hanzhe Zhang. 2021. “Preference Evolution in Different Matching Markets.” *European Economic Review* 137:103804.