Evolutionary stability of equal sharing

Jiabin Wu* Hanzhe Zhang†

April 2, 2025

Abstract

This paper offers an evolutionary explanation for the persistence of the equal sharing norm, which promotes fairness but is often viewed as less efficient than market-based alternatives. We present a model in which agents, differentiated by productive types, form partnerships endogenously. While equal sharing may lead to less efficient short-term outcomes, it fosters the growth of productive types, which ensures long-term efficiency. Unequal sharing norms, by contrast, generate inefficiencies and diminish over time. Our findings rationalize the coexistence of equal sharing and market norms.

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

This paper provides an evolutionary justification for the equal sharing norm. The equal sharing norm, which advocates for the equitable distribution of resources among individuals, 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*, remains widely adopted by modern law firms (Smith, 2001). Equal sharing of loyalties among band members is a common practice, even among the most renowned bands (Polcz, 2023). Kibbutz is a vibrant Israeli community that promotes equality of 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 inflexible. 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, where the output of each partnership depends on the productive types 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, as in a transferable utility (TU) matching market. Second, the **equal sharing norm** requires matched pairs to divide output equally. Third, we consider **unequal sharing norms**, such as those that allocate a larger share to the more senior agent. Unlike the market sharing norm, equal and unequal sharing norms lack flexibility, causing the matching mechanism to operate as a non-transferable utility (NTU) matching market.

Our model assumes that, at the behavioral level, agents within a population engage in a repeated re-matching process, which quickly converges to a stable matching. At the type level, the distribution of types evolve more slowly, driven by the payoffs earned through matching, and converge to a stable distribution. Finally, at the societal level, sharing norms evolve at the slowest rate, shaped by competition among multiple populations with different sharing norms, based on their societal performances.

We find that while the equal sharing norm may not produce as efficient stable matchings as the market sharing norm at the behavioral level, both lead to efficiency at the level of type evolution. This means they ultimately favor the selection of agent types that generate the highest partnership outputs. In contrast, unequal sharing norms introduce inefficiencies at the type evolution stage. As a result, over time, unequal sharing norms are outcompeted. However, both the market sharing norm and the equal sharing norm persist and can coexist in the long run.

To the best of our knowledge, no paper has discussed a *three-tier evolutionary model* with endogenous matching. Previous papers allow the evolution of productive types, given a rapid matching process and a fixed sharing rule over time (e.g., Wu and Zhang, 2021; Hiller, Wu, and Zhang, 2023; Wang and Wu, 2025).

In our model, the partnership formation process is modeled as a one-sided pairwise matching markets (i.e., roommate markets) 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. As a result, we provide a companion paper (Wang, Wu, and Zhang, 2024) that states in more detail and proves more rigorously these results, which may serve as a reference for subsequent papers that model the evolution of behavior/productive types/preferences in one-sided matching markets.

The only other work that explores evolutionary justifications for the equal sharing norm is 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 replicator dynamics, a homogeneous population of fair-minded agents and 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 diverges from Skyrms (2014) in several key aspects. First, agents in our model exhibit heterogeneity in their productive types, rather than heterogeneity in demand rules. Second, partnerships in our model emerge endogenously, as opposed to being formed through random encounters. Third, the selection of sharing norms occurs at the societal level, rather than the individual level, and is driven by the economic performance associated with each sharing norm.

2 The Market Sharing Norm

2.1 Population and Matching

Consider a unit mass of agents. Each agent is characterized by a trait $i \in N = \{1, 2, ..., n\}$. Let $x_i \in [0, 1]$ denote the mass of agents who carry trait i, 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. 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}.[1]

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 $f(i, j)$, where $f : N^2 \to \mathbb{R}_{++}$ is 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 notion of individual matching.

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 markets with a continuum of agents to characterize the stable outcome of the partnership formation process (Wang, Wu, and Zhang, 2024).

Definition 1 *A **stable outcome** is described by a **stable aggregate 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 $i, j \in N$.*

We discuss 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

[1] In our 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.

agent, all agents are matched in any stable aggregate 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.

Just as in a two-sided matching market (Gretsky, Ostroy, and Zame, 1992), an aggregate matching is stable if and only if it is socially efficient (Wang, Wu, and Zhang, 2024). Hence, we can characterize a stable aggregate 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 aggregate matching. We will assume that the stable aggregate matching is always unique in the subsequent analysis.[2] Let $W^*(x)$ be the collection of stable payoff vectors $w^*(x)$ that support the unique stable aggregate matching $\mu^*(x)$.

Example 1 *Consider two types, $N=\{1,2\}$. First, suppose $2f(1,2)<f(1,1)+f(2,2)$. The unique stable outcome involves positive assortative matching—$\mu^*_{ii}=x_i/2$ 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. When $x_1<1/2$, the unique stable aggregate 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 aggregate 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 aggregate 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.[3]

[2] A sufficient condition to ensure unique stable outcome is $f(i,j)+f(i',j')\neq f(i,i')+f(j,j')$ for any $i,i',j,j'\in N$ such that $\{i,i'\}\cap\{j,j'\}=\emptyset$.

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

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 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 an n-dimensional vector in which each of the n entries equals $\overline{w}(x)$. The replicator dynamic is given by a differential inclusion:[4]

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

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 equation (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, suppose 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.

In Appendix A, we prove the existence of a solution trajectory for RDI. Note that the uniqueness of the solution trajectory is not guanranteed.

2.3 Analysis

In this section, we study convergence and stability of dynamics. Convergence informs us where the dynamic ends up at from all initial conditions, while stability determines whether convergent states of the population are robust against perturbations. See Appendix A for the definitions of convergence, stability, and Theorem 2 and 3 that we use to prove stability and convergence based on the Lyapunov method.

We will focus on the case that there exists a unique match that leads to the highest payoff. There are two subcases. 1) Suppose there exists a unique i such that $f(i,i) = \max_{k,l} f(k,l)$. In this case, we define $L^{RDI}(x) = -\log x_i$. 2) There exist i and j such that $i \neq j$ and $f(i,j) = \max_{k,l} f(k,l)$. In this case, we define $L^{RDI}(x) = -(\log(2x_i) + \log(2x_j))$. Proposition 1 will use $L^{RDI}(x)$ as the Lyapunov function for (RDI).

[4] 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 function.

Proposition 1 ***(i)*** *Suppose there exists a unique i such that $f(i,i) = \max_{k,l} f(k,l)$. Then the population state $x^* = e_i$, the ith standard basis vector, is globally asymptotically stable 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 such that $i \neq j$ and $f(i,j) = \max_{k,l} f(k,l)$. Then the population state x^*, such that $x_i^* = x_j^* = 0.5$, is globally asymptotically stable 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.*

Proof: 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 $v = x \cdot [w^*(x) - \bar{w}^*(x)]$, we have

$$\begin{aligned} \frac{\partial L^{RDI}}{\partial v}(x) &= \nabla L^{RDI}(x)'v \\ &= -1/x_i * x_i(w_i^*(x) - \bar{w}^*(x)) \\ &= -(w_i^*(x) - \bar{w}^*(x)). \end{aligned}$$

By the definition of stable matching, $w_i^*(x) \geq f(i,i)/2$. By the fact that $f(i,i) = \max_{k,l} f(k,l)$, $f(i,i)/2 > \bar{w}^*(x)$ 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 $w_i^*(e_i) = \bar{w}^*(e_i) = f(1,1)/2$. Hence, $V^{RDI}(e^1) = \{\mathbf{0}\}$. Also, $L^{-1}(0) = \{e_i\}$. By Theorem 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 uniquely 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 $v = x \cdot [w^*(x) - \bar{w}^*(x)]$, we have

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

By the definition of stable matching, $w_i^*(x) + w_j^*(x) \geq f(i,j)$. By the fact that $f(i,j) = \max_{k,l} f(k,l)$, $f(i,j)/2 > \bar{w}^*(x)$. 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 Theorem 2 and 3, x^* is the globally asymptotically stable state with respect to $X\backslash\{x|x_i = 0$ or $x_j = 0$ or both$\}$. In addition, $f(i,j) = \max_{k,l} f(k,l)$ implies that x^* is the uniquely efficient state. ∎

Proposition 1 shows that the populations states that maximize the average payoff in stable matching are asymptotically stable.

Example 2 *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, which captures complementarity between the two traits. 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, which captures substitutability between the two traits. There is always a unique stable matching with only own-trait matching (positive assortative matching).*

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 $x_1 < 1/2$, $w_1(x) = f(1,2) - f(2,2)/2 > w_2(x) = f(1,2) - f(2,2)$. When $x_1 > 1/2$ $w_(x) = f(1,1)/2 < w_2(x) = f(1,2) - f(1,1)/2$. Hence, for (RDI), $\dot{x}_i > 0$ when $0 < x_1 < 1/2$; $\dot{x}_i < 0$ when $x_1 > 1/2$, which guarantees the global asymptotic stability of $x_1 = 1/2$ in $(0,1)$.*

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

To summarize, while supermodularity guarantees 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 type-i agent's payoff from her match with a type-j agent. $u_{ij} + u_{ji} = f(i,j)$. We adopt the stability concept in two-sided NTU (non-transferable 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 2 *A **stable aggregate 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 $\mu^*(ik) > 0, \mu^*(jl) > 0$.*

Carmona and Laohakunakorn (2024) prove the existence of stable aggregate matching.

For x and $\mu^*(x)$, let $u^*_i(x)$ denote the average payoff earned by all type-i agent. Let $u^*(x) = (u^*_1(x), u^*_2(x), ..., u^*_n(x)) \in \mathbb{R}^n$. Let $\bar{u}^*(x)$ denote the average payoff in the entire population. Let $\bar{\mathbf{u}}^*(x) = (\bar{u}^*(x), \bar{u}^*(x), ..., \bar{u}^*(x)) \in \mathbb{R}^n$. The evolution of traits is still dictated by the replicator dynamic. But now it is given by a differential equation:

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

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

3.1 Equal Sharing Norm

Suppose two agents in a pair equally split the payoff they jointly produce. That is, $u_{ij} = f(i,j)/2$.

In the remaining analysis, we will assume that $f(i,j) \neq f(i,l)$, for any i and any $j \neq l$ (see footnote 2). Without loss of generality, assume $f(1,1) > f(i,i)$ for any $i \neq 1$.

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

Proposition 2 ***(i)*** *Suppose $f(1,1) = \max_{k,l} f(k,l)$. Then $x = e_1$, the 1^{th} standard basis vector, is the 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.*

Proof: 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 i-type agents must match among themselves and $F_1(x) = u_{11} = f(1,1)/2$ 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 $x_1^t > 0$,

$$
\begin{aligned}
dL^{RDE}(x^t)/dt &= \bigtriangledown L^{RDE}(x^t)'\dot{x}^t \\
&= -1/x_1^t * x_1^t(F_1(x^t) - \overline{F}(x^t)) \\
&= \overline{F}(x^t) - F_1(x^t) \\
&< 0 \text{ as long as } x^t \neq e_1.
\end{aligned}
$$

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 types 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 type-i agents are matched with type-j agents, while the remaining type-j agents match otherwise. In this case, $F_i(x^0) = u_{ij} = f(i,j)/2$, which is the largest payoff an agent can get.

Let $r(t) = x_i^t/x_j^t$. Observe that under the replicator dynamic, $\frac{dr(t)}{dt} = \frac{x_i^t}{x_j^t}[F_i(x^t) - F_j(x^t)] > 0$ as long as $x_i^t < x_j^t$ so that $F_i(x^t) - F_j(x^t) > 0$. Also, once $x_i^t = x_j^t$, $r(t) = 1$ and $F_i(x^t) = F_j(x^t) = f(i,j)/2$ because all type-i agents are exactly matched with all type-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,

$$
\begin{aligned}
dL^{RDE}(x^t)/dt &= \bigtriangledown L^{RDE}(x^t)'\dot{x}^t \\
&= -1/2x_i^t * x_i^t(F_i(x^t) - \overline{F}(x^t)) - 1/2x_j^t * x_j^t(F_j(x^t) - \overline{F}(x^t)) \\
&= \overline{F}(x^t) - (F_i(x^t) + F_j(x^t))/2 \\
&= \overline{F}(x^t) - f(i,j)/2 \\
&< 0 \text{ as long as } x^t \neq x^*.
\end{aligned}
$$

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 $X \backslash \{x | x_i = 0 \text{ or } x_j = 0 \text{ or both}\}$. ■

One may observe 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 types (including potentially the same type) are most productive when matched, these types 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 types 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 types exceeds the population average.

Example 3 *Suppose there are three types.* $f(1,1) > f(1,2)$, $f(1,1) > f(1,3)$, $f(2,2) > f(1,2)$, $f(2,2) < f(2,3)$, $f(3,3) < f(1,3)$, $f(3,3) > f(2,3)$. *This payoff structure leads to PAM. Figure 1 provides a graphic illustration of* (RDE) *given* $f(1,1) = 6$, $f(1,2) = 1$, $f(1,3) = 5$, $f(2,2) = 2$, $f(2,3) = 3$, $f(3,3) = 4$.

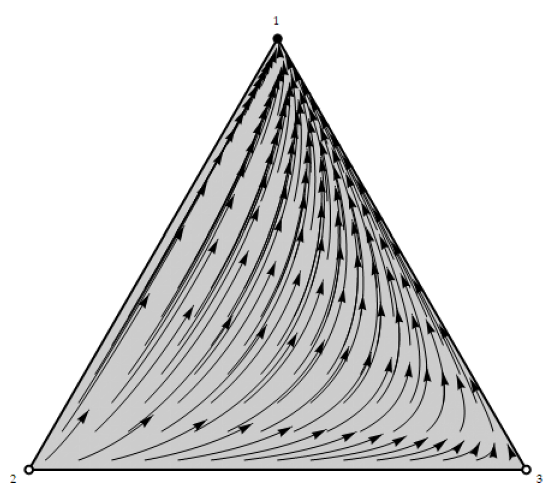

Figure 1: Replicator dynamic under PAM.

Example 4 *Suppose there are three types.* $f(1,1) < f(1,2) < f(1,3)$, $f(2,2) < f(1,2) < f(2,3)$, $f(3,3) < f(1,3) < f(2,3)$. *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* $\mu_{1,2} = 0.3, \mu_{2,3} = 0.2$. *Then* $F_1(x^0) =$

$f(1,2)/2$, $F_2(x^0) = 0.6 * (f(1,2)/2) + 0.4(f(2,3)/2)$ and $F_3(x^0) = f(2,3)/2$. The average payoff $\overline{F}(x^0) = 0.6 * (f(1,2)/2) + 0.4(f(2,3)/2)$. Hence, under the replicator dynamic, $\dot{x}_1 = x_1^0(F_1(x^0) - \overline{F}(x)) < 0$, $\dot{x}_2 = x_2^0(F_2(x^0) - \overline{F}(x)) = 0$, $\dot{x}_3 = x_3^0(F_3(x^0) - \overline{F}(x)) > 0$. Hence, x_1 shrinks, x_3 grows and x_2 stagnates and $\lim_{t\to\infty} x^t = (0, 0.5, 0.5)$.

Figure 2 provides a graphic illustration of (RDE) given $f(1,1) = 1$, $f(1,2) = 3$, $f(1,3) = 5$, $f(2,2) = 2$, $f(2,3) = 6$, $f(3,3) = 4$.

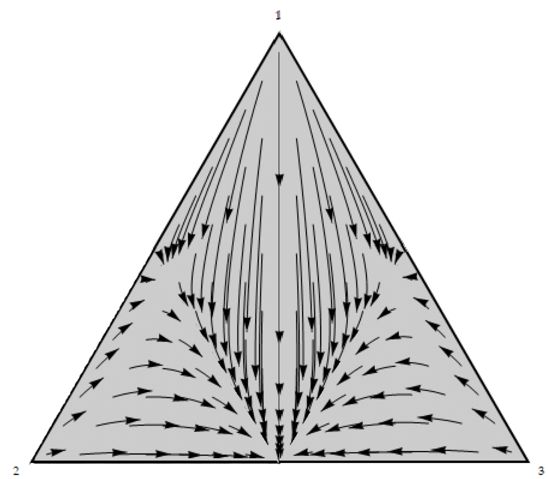

Figure 2: Replicator dynamic under Non-PAM.

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

3.2 Unequal Sharing Norms

Suppose the society has a universal sharing norm that features an unequal division of joint production. Let $\alpha_{i,j} \in (0,1)$ denote the share of joint production a type-i agent gets when matched with a type-j agent.

Assume that the universally unequal sharing norm takes the following form: $\alpha_{i,i} = 0.5$ for any i and $\alpha_{i,j} > 0.5$ if $i < j$. The society is hierarchical in this case, where types represent seniority and a relatively more senior agent can get a larger share when matched with a relatively more junior agent. Assume $f(k,k) = \max_i f(i,i)$, indicating that type-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 most senior type is also the most productive in own-type matching.

For illustration, consider the following examples:

Example 5 *Suppose there are two types 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, as the society can achieve a higher average payoff at the state* $(0.5, 0.5)$.

Example 6 *Suppose there are two types 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 type-1 and type-2 agents. When* $x_1 \leq 0.5$,

$$F_1(x) = u_{12} = 6 > F_2(x) = \frac{x_1}{x_2}u_{21} + \frac{x_2 - x_1}{x_2}u_{22} = 4\frac{x_1}{x_2} + 3.5\frac{x_2 - x_1}{x_2}.$$

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

$$F_1(x) = \frac{x_1 - x_2}{x_1}u_{11} + \frac{x_2}{x_1}u_{12} = 3\frac{x_1 - x_2}{x_1} + 6\frac{x_2}{x_1}, \ F_2(x) = u_{21} = 4.$$

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

Proof: 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 $i \neq 1$. ■

4 The Evolution of Sharing Norms

Suppose that on top of the evolution of traits, there is another layer of evolution on sharing norms. Let s denote a sharing norm and let $\mathbf{S}$ denote the space for sharing norms. There are many populations, and each is equipped with a different sharing norm. Let $x^*(s)$ denote a stable state in the evolution of types in the population with sharing norm s and let $X^*(s)$ denote the set of stable states.

Definition 3 *A sharing norm s is evolutionary stable s' if for every payoff structure f, $\overline{F}(x^*(s)) \geq \overline{F}((x^*(s'))$ for any $x^*(s) \in X^*(s)$ and $x^*(s') \in X^*(s')$; and for some payoff structure f, $\overline{F}(x^*(s)) > \overline{F}((x^*(s'))$ for some $x^*(s) \in X^*(s)$ and $x^*(s') \in X^*(s')$.*

We are essentially considering a three-speed dynamic model. At the behavior level, agents within each population engage in a repeated re-matching process, which quickly converges to a stable matching. At the trait level, the distribution of traits evolves relatively more slowly and converges to a stable distribution. At the societal sharing norm level, the sharing norms evolve the slowest. According to our previous analysis, the equal sharing norm and the market sharing norm are robust in selection on the society level, as they always induce the efficient distribution of traits.

Proposition 4 *The equal sharing norm and the market sharing norm are both evolutionarily stable against any non-equal sharing norms.*

Proof: Directly implied by Proposition 1, 2 and 3. ■

This explains why equal sharing norms would still survive in markets.

5 Conclusions

This paper offers an evolutionary perspective on the resilience of the equal sharing norm, challenging the assumption that market-based approaches are inherently superior. Through a three-tier evolutionary framework, we demonstrate how equal sharing, despite its rigidity, fosters efficient type evolution, promoting 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.

By extending the existing literature to incorporate endogenous partnership formation and the gradual evolution of societal norms, our model highlights the interplay between economic performance and norm selection. The results underscore the importance of considering multi-level evolutionary processes when analyzing social and economic institutions.

References

Abramitzky, Ran. 2008. "The Limits of Equality: Insights from the Israeli Kibbutz." *The 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.

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.

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

Güth, W. and M. 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 U. Witt. University of Michigan Press, 23–34.

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

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, T.A. 2001. "Equality, Evolution and Partnership Law." *Journal of Bioeconomics* 3:99–121.

Taylor, P.P. and L. 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.

Wang, Ziwei, Jiabin Wu, and Hanzhe Zhang. 2024. "The Nonatomic Roommate Problem." Mimeo.

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

A Appendix on Dynamic Properties

Theorem 1 *(Sandholm, 2010, Theorem 6.A.1) Consider a differential inclusion $\dot{x} \in V(x)$. If $V(x)$ is nonempty, bounded, convex-valued and upper-hemicontinuous, then for any initial condition $\xi \in \mathbb{R}^n$, there exists a Carathéodory solution $\{x_t\}_{t\in[0,T]}$ 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 has a solution trajectory in the interior of X.

Lemma 1 *The correspondence V^{RDI} is nonempty-valued, bounded, convex-valued, and upper-hemicontinuous in the interior of X.*

The limiting behavior of deterministic dynamics can be characterized as follows.

Definition 4 *The ω-limit set $\omega(\xi)$ is the set of all points that the solution trajectory $\{x_t\}_{t\geq 0}$ starting from $x_0 = \xi$ approaches arbitrarily closely infinitely often:*

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

The set of all ω-limit points of all solution trajectories is given by $\Omega = \cup_{\xi \in X}\omega(\xi)$. It captures the notion of recurrence of the dynamic.

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 5 *A proper subset of X, A, is **Lyapunov stable** if for every neighborhood O of A, there exists a neighborhood O' of A such that every solution $\{x_t\}_{t\geq 0}$ that starts in O' is contained in O, that is, $x_0 \in O'$ implies that $x_t \in O$ for all $t \geq 0$.*

Intuitively, Lyapunov stability requires that any displacement from A does not lead the process to go "far" from A at any point in time. If A is not Lyapunov stable, we call it ***unstable***.

Definition 6 *A is **attracting** if there is a neighborhood Y of A such that every solution that starts in Y converges to A, that is, $\xi \in Y$ implies $\omega(\xi) \in A$. A is **globally attracting** if it is attracting with $Y = X$.*

The set of points $\xi \in X$ such that starting at ξ, $\lim_{t\to\infty} x_t \in A$ is called the *basin of attraction* of A. Intuitively, attracting requires that given any displacement from A, the process returns to A in the limit.

Definition 7 *A is **asymptotically stable** if it is Lyapunov stable and attracting. A is **globally asymptotically stable** if it is Lyapunov stable and globally attracting.*

Intuitively, asymptotic stability requires that given any displacement from A, the process never travels 'far' from A and returns to A in the limit.

In what follows, we provide the theorems we will use 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) $\left[\boldsymbol{0} \notin V^{RDI}(x) \implies \frac{\partial L}{\partial v}(x) < 0\right]$ for all $v \in V^{RDI}(x)$ and $x \in Y$. Then, for all solutions $\{x_t\}$ of (RDI) with $x_0 \in Y$, $\omega(\{x_t\}) \subseteq \{x \in Y : \boldsymbol{0} \in V(x)\}$.*

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).

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).