Economics Working Paper

2011-12

Do Contracts Help?
A Team Formation Perspective

Norovsambuu Tumennasan

Department of Economis and Business
Business and Social Sciences
Aarhus University
Bartholins Allé 10, Building 1322
DK-8000 Aarhus C - Denmark
Phone +45 8942 1610
Mail: oekonomi@econ.au.dk
Web: www.econ.au.dk

Do Contracts Help? A Team Formation Perspective*

Norovsambuu Tumennasan[†‡]
School of Economics and Management, Aarhus University

First Version: July 2008
This Version: September 2011

Abstract

Economists perceive moral hazard as an undesirable problem because it undermines efficiency. Carefully designed contracts can mitigate the moral hazard problem, but this assumes that a team is already formed. This paper demonstrates that these contracts are sometimes the reason why teams do not form. Formally, we study the team formation problem in which the agents' efforts are not verifiable and the size of teams does not exceed quota r. We show that if the team members can make only balanced transfers, then moral hazard affects stability *adversely*. However, if the team members cannot make transfers, then moral hazard affects stability *positively* in a large class of games. For example, a stable team structure exists if teams produce public goods or if the quota is two. However, these existence results no longer hold if efforts are verifiable.

Keywords: team formation, hedonic game, moral hazard, assortative partition

JEL classification: C71; C78

---

*This paper is a chapter of my Ph.D. dissertation at Brown University and was circulated as "A Model of Coalition Formation with Production under Moral Hazard."

[†]Email: ntumennasan@econ.au.dk.

[‡]I am indebted to my advisors Roberto Serrano, Geoffroy de Clippel, and Sergio Turner for their continual advice and encouragement. Special thanks to David Perez-Castrillo whose comments helped to shape this version of the paper. I also thank Debipriya Chatterjee, John Kennes, Alexander Koch, Daniel Monte, Petros Milionis and the seminar participants at Aarhus, Brown and Maastricht Universities for their useful comments and suggestions. All errors are my own.

1 Introduction

Economists have long recognized the importance of the moral hazard problem: a Google Scholar search returns more than 3000 articles with titles containing the term "moral hazard." This vast literature offers two rather robust conclusions: (1) moral hazard undermines efficiency and (2) carefully designed contracts can mitigate the moral hazard problem. A typical article in the "moral hazard" literature considers a team (possibly 1 agent team) and then searches for possible ways to improve efficiency under moral hazard. However, the literature is silent on how these teams came to existence. This paper contributes in filling this gap. In other words, we investigate what team structure — a collection of teams — would emerge if the agents had freedom to form a team. Formally, we study the team formation problem when the efforts of the agents are not verifiable.

Once the agents can form a team, the most important issue is stability: because the agents endogenously form teams leading to a certain team structure, no group of agents should be able to improve themselves by forming a new team. Hence, efficiency is of secondary importance in our team formation problem.

In this paper, the agents with different productivities endogenously form teams that produce local goods. The production outcome (deterministically) depends on the total productivity weighted efforts of the team members, i.e., the efforts of different agents are perfect substitutes. To accommodate the roommate problems, we impose an exogenous quota or limit r on the team size.

When efforts are not verifiable, a credibility issue arises, i.e., teams cannot supply certain effort profiles because individual members can shirk without being identified. What effort profiles are credible depends on the transfer the team members can make amongst themselves. We consider two cases: (1) the team members can make *balanced* transfers only or (2) the team members cannot make any transfers.

Our first main result shows that if the team members can make balanced transfers only, then moral hazard affects stability negatively. This result is directly based on the results of Hölmström (1982) and Legros and Matthews (1993). Hölmström (1982) shows that, through balanced transfers, teams cannot achieve the efficient level of welfare when efforts are not verifiable. On the other hand, Legros and Matthews (1993) demonstrate that, through balanced transfers, teams can approximate their efficient welfare to any degree. Therefore, the utility possibility set for any team is open, implying that any two or more agents cannot agree on a balanced transfer as there is always another one which improves them. Consequently, if there is any stable structure, then each agent works alone. This result is robust, as it does not rely on the assumption that the efforts of different agents are perfect substitutes. In addition, our first main result implies that the core — a fundamental concept in cooperative game theory — is empty if efforts are not verifiable and the agents can make only balanced transfers.

Our second main result shows that if the team members cannot make transfers at all, then moral hazard positively affects stability in a large class of games. For example, if teams produce public goods or if the quota is 2 (the standard roommate problem), then a stable team structure always exists when efforts are not verifiable. However, this is not the case if efforts are verifiable. Hence, the inability to make transfers could completely reverse the effect of moral hazard on stability. In general, the combination of moral hazard and the inability to make transfers leads to the worst efficiency loss. Yet in terms of stability, this combination is desirable.

A key reason for our second main result is that each team has a unique credible effort profile once team members cannot make transfers. Hence, whether a team can block an already formed team structure depends on its credible effort profile. To prove our existence results, we construct a simple algorithm: the most productive agent chooses her team, and then the most productive *unmatched* agents chooses her team out of the *unmatched* agents and so on. The proofs use the finding that, the team an agent prefers, out of teams of equal size, is that of the most productive agents. In addition, if teams produce public goods, recruiting an additional agent is always beneficial to each team member. Consequently, the preference of the most productive agent perfectly aligns with the preferences of the agents she chooses in her team. Similarly, if the quota is 2, again the preferences of the team members (possibly one) the most productive agent chooses in her team are also aligned. Hence, no team can break down the team structure resulted from the canonical algorithm. If the quota exceeds 2 and the local good is not public, then some of the members of the team the most productive agent chooses might prefer a smaller team. By exploiting this disagreement, we construct an example with no stable team structure for this case.

We also show that if efforts are not verifiable and the team members cannot make transfers, then, in a large class of games, the stable team structures are assortative by productivity, i.e., the more productive agents form a team while the less productive agents form a team. This is the case, for example, if teams produce public goods or if the quota is two. Assortativeness is an interesting property which seems to be observed in life. However, one cannot generalize this result to non-public goods when the quota exceeds 2. Hence, the assortativeness property seems to be restricted only to the public good or the quota 2 cases.

The team formation problem under moral hazard has not been explored to my knowledge. Dam and Prez-Castrillo (2006) consider the principal agent model in a two sided matching markets setting: each player is either a principal or an agent and each principal matches with an agent. We do not have this restriction in our paper. The "moral hazard in teams" literature started by Hölmström (1982) studies whether some contract can deliver the efficient outcomes under moral hazard. This paper, on the other hand, studies whether teams endogenously form under moral hazard.

Bogomolnaia and Jackson (2002) and Banerjee et al. (2001) study team formation for hedonic games. In hedonic games, the players' payoffs from a given team depend on the identities the team members. If the team members cannot make transfers, then our team formation game under moral hazard is a hedonic game as there is a unique Nash equilibrium. The above mentioned papers provide sufficient conditions for the existence of stable partitions. In fact, our team formation problem under moral hazard satisfies the *(weak) top team* property defined by Banerjee et al. (2001) if teams produce public goods or if the quota is 2. A well studied subclass of hedonic games is the roommate problem[1] which is a team formation problem with a quota of two. When the quota is 2, our team formation problem under moral hazard satisfies the *no odd rings* condition defined by Chung (2000).

The paper is organized as follows: we build the model in section 2, and study team formation in the absence of moral hazard in section 3. We investigate team formation in the presence of moral hazard in section 4. Lastly, we conclude in section 5.

2 Model

Let $N = \{1, ..., n\}$ be the set of players and let $S \subseteq N$ stand for a team. The model consists of two periods: team formation and production.

In the team formation period, the players form teams of size r or less. We shall refer to r as the quota. Various factors justify the quota;physical constraints (the size of the laboratory), coordination issues (too hard to coordinate when there are more than r players), or social norm. The use of quota allows us to accommodate the roommate problem in our model.

Each player can be a member of only one team. Thus, we define the concept of r-partition to model the team formation period: r-partition $\Pi_r = \{S_1, \cdots, S_m\}$ is a partition of N such that, for any $S \in \Pi_r$, $|S| \leq r$. As we deal with r-partitions only, we say "partition" instead of "r-partition." In addition, the notation $\Pi_r(i)$ denotes the team which includes player i under partition Π_r.

In the production period, each team produces local goods that benefit only its members.[2] Player i is endowed with a productivity $\lambda_i \in \mathbb{R}_{++}$ and supplies an effort e_i in the production period. Each team's production function $f(\cdot)$ depends on the team's total productivity weighted effort (*weighted effort* from here on), i.e., $f\left(\sum_{i \in S} \lambda_i e_i\right)$. Consequently, the efforts of different players are perfect

[1] For more information see Gale and Shapley (1962), Tan (1991), and Chung (2000).

[2] Local public goods are sometimes known as club goods. For more information, see Buchanan (1965) and Sandler and Tschirhart (1980).

substitutes, which is a somewhat restrictive assumption. There is another interpretation to our model: in the production period, each team produces *one* unit of local goods and the quality of local goods f depends on the weighted effort. We impose the following restrictions on the production function.

Assumption 1. *The production function f is twice differentiable, strictly increasing, and concave, i.e., $f' > 0$ and $f'' \leq 0$.*

Let $\rho : \mathbb{N} \to [0, 1]$ be the "team" credit function; if team S produces x amount of local goods, then its each member can consume $\rho(|S|)\, x$ amount of local goods. The credit function allows us to analyze a broad range of goods. For example, the credit function with $\rho(\cdot) = 1$ corresponds to public goods while the one with $\rho(s) = 1/s$ for all $s \in \mathbb{N}$ corresponds to private goods. In addition to these two well-known cases, there are some intermediate cases of interest. For example, suppose a researcher is being considered for tenure. Her research papers are evaluated; a coauthored paper is valued less than 1 but not less than 1/2 single-authored paper assuming that the quality of the papers are the same. We can model this case easily thanks to the credit function.

The credit function depends on the size of a team, but not on the team itself. Furthermore, the credit function is not individual specific, which is somewhat restrictive. We normalize $\rho(1) = 1$ and assume that the credit function is a non-increasing function, i.e., the members of a smaller team receive more credit than the ones of a bigger team.

The quota can be modeled using the credit function; by setting $\rho(s) = 0$ for $s > r$, one effectively eliminates the formation of teams with size greater than r. However, for the interpretational purposes we use the quota explicitly in our analyses.

The team members can transfer money among themselves. A transfer scheme for team S specifies how much transfer one receives based on the production outcome. We restrict our attention to the *balanced* transfer schemes, i.e., the total transfer for each team is 0. Formally, a transfer scheme for team S is a function $\phi^S : \mathbb{R}_+ \to \mathbb{R}^{|S|}$ satisfying $\sum_{i \in S} \phi_i^S(x) = 0$ for all $x \in \mathbb{R}_+$. Observe that the 0 transfer schemes indicate the cases in which the team members cannot/do not make transfers.

The transfer schemes do not depend on the efforts. This might seem too restrictive for the verifiable effort case. However, this is not the case for the following reason: when efforts are verifiable, any transfer scheme dependent on the efforts can be rewritten in terms of the production outcome as the team members can contract on efforts and the production function is deterministic.

The players's utility depends on the consumption of local good (x_i) and effort (e_i) as follows:

$$U_i(x_i, e_i) = x_i - c(e_i)$$

where $c : \mathbb{R}_+ \to \mathbb{R}_+$ is the cost of effort.

Assumption 2. *The cost function $c(\cdot)$ is twice differentiable, strictly increasing, strictly convex, $c(0) = 0$ and $c'(0) = 0$. In addition, there exists $\bar{e}$ such that $\lambda_i f'(\lambda_i \bar{e}) < c'(\bar{e})$ for all $i \in N$.*

Thanks to assumption 2, each player exerts a finite effort if one stays unmatched.

If the members of team S supply an effort profile $e^S = (e_i^S)_{i \in S}$ and make transfers according to ϕ^S, the utility of player $i \in S$ is:

$$U_i(S, \phi^S, e^S) = \rho\left(|S|\right) f\left(\sum_{j \in S} \lambda_j e_j^S\right) + \phi_i^S\left(f\left(\sum_{j \in S} \lambda_j e_j^S\right)\right) - c(e_i^S) \tag{1}$$

Finally, suppose the players form teams according to a partition Π_r and each team $S \in \Pi_r$ exerts e^S and make transfers according to ϕ^S. Let $e^{\Pi_r} = \left(e^S\right)_{S \in \Pi_r}$ and $\phi^{\Pi_r} = \left(\phi^S\right)_{S \in \Pi_r}$. Then player i's utility from $\left(\Pi_r, \phi^{\Pi_r}, e^{\Pi_r}\right)$ is $U_i\left(\Pi_r, \phi^{\Pi_r}, e^{\Pi_r}\right) = U_i\left(\Pi_r(i), \phi^{\Pi_r(i)}, e^{\Pi_r(i)}\right)$. If the transfer scheme is 0, we simplify the notations by using $U_i(S, e^S)$ for $U_i(S, 0, e^S)$ and $U_i\left(\Pi_r, e^{\Pi_r}\right)$ for $U_i\left(\Pi_r, 0, e^{\Pi_r}\right)$.

Before analyzing the team formation problem, let us introduce the notion of assortative partition which is used in the matching literature: under an assortative partition, the more productive players form a team while the less productive players form a team. Formally, partition Π_r is assortative (by productivity) if, for any two distinct teams $S, T \in \Pi_r$, either $\min_{i \in S} \lambda_i \geq \max_{i \in T} \lambda_i$ or $\min_{i \in T} \lambda_i \geq \max_{i \in S} \lambda_i$.

3 Team Formation in the Absence of Moral Hazard

To study the effects of moral hazard on team formation, this section investigates team formation under verifiable efforts.

When efforts are verifiable, teams can contract on efforts. Suppose that the players are about to form teams which would lead to partition Π_r. Each team fixes its transfer scheme and effort profile. Hence, the players know their utility if the current plan is implemented. Now the players start looking for better options: if some players can improve themselves over the status quo by forming a team and specifying the corresponding transfer scheme and effort profile, they will abandon the current plan. If not, the current plan is implemented. This is the logic behind the definition of blocking and stable partition. We consider two notions of stability depending on whether the team members can make transfers.

Definition 1. *Team S blocks $(\Pi_r, \phi^{\Pi_r}, e^{\Pi_r})$ with (ϕ^S, e^S) if*

$$U_i(S, \phi^S, e^S) > U_i\left(\Pi_r, \phi^{\Pi_r(i)}, e^{\Pi_r(i)}\right) \text{ for all } i \in S$$

1. *A partition Π_r is stable if, for some ϕ^{Π_r} and e^{Π_r}, no team S blocks $(\Pi_r, \phi^{\Pi_r}, e^{\Pi_r})$ with any ϕ^S and e^S.*

2. *A partition Π_r is stable under 0 transfer schemes if, for some e^{Π_r}, no team S blocks $(\Pi_r, 0, e^{\Pi_r})$ with any $(0, e^S)$.*

Before investigating the existence of stable partitions, let us consider the utilities that each team can achieve using transfer schemes when efforts are verifiable. This is instrumental in our analysis of team formation under non-verifiable efforts, as we will see in the next section.

As the players' utilities are linear in local goods, each team's welfare (the sum of utilities) for a given effort profile e^S is

$$W^S(e^S) = |S|\rho(|S|)f\left(\sum_{i\in S}\lambda_i e_i^S\right) - \sum_{i\in S}c(e_i^S).$$

An effort profile is efficient for S if it maximizes the welfare for S. Our assumptions on the cost and production functions guarantee the existence of a unique efficient effort profile for each team. Let the welfare corresponding to the efficient effort profile be $\bar{W}^S$. Clearly, team S can achieve an utility vector $u^S \in \mathbb{R}_+^{|S|}$ if and only if $\sum_{i\in S} u_i^S \le \bar{W}^S$. In fact, team S can achieve any utility vector $u^S \in \mathbb{R}_+^{|S|}$ with $\sum_{i\in S} u_i^S = \bar{W}^S$ through a balanced transfer scheme. This means that the utility possibility set and the pareto frontier for coalition S are $V^S = \{u^S \in \mathbb{R}_+^{|S|} : \sum_{i\in S} u_i^S \le \bar{W}^S\}$ and $\partial V^S = \{u^S \in \mathbb{R}_+^{|S|} : \sum_{i\in S} u_i^S = \bar{W}^S\}$, respectively.

Now, we are ready to study the existence of stable partitions. Proposition 1 states that if the teams produce public goods and there is no quota restriction, then a stable partition exists. The main intuition for this result is that the public nature of the local good eliminates the penalty for forming a bigger team. In addition, because there is no quota restriction, the players can form one big team.

Proposition 1. *If teams produce public goods and there is no quota restriction ($r = n$), then*

1. *$\Pi = \{N\}$ is the unique stable partition*

2. *$\Pi = \{N\}$ is the unique stable partition under 0 transfer schemes*

Proof. See Appendix. □

The main idea of the proof is to transform our team formation problem into cooperative game defined by Scarf (1967) and then show that this game is balanced. The core for balanced games is not empty implying that the only stable partition is $\Pi = \{N\}$. The non-emptiness of the core is not surprising since similar results were established in public good economies which are somewhat different than ours.[3]

If the quota is binding ($r < n$) or if the credit function is strictly decreasing, then there might not exist a stable partition. We illustrate this point in the next example.

Example 1. *Let $r = 2$ and $N = \{1, 2, 3\}$. If $\rho(\cdot) = 1$ and $\lambda_1 = \lambda_2 = \lambda_3 > 0$, then*

1. *there does not exist stable partition*
2. *there does not exist stable partition under 0 transfer schemes.*

Proof. Let $e^* = \arg\max_{e_i} U_i(\{i\}, e_i)$. Then if player i works alone, then her utility is $U_i(\{i\}, e^*)$.

1. Suppose a stable partition Π exists. Thanks to proposition 1, $\Pi \neq \{\{1\}, \{2\}, \{3\}\}$. Hence, without loss of generality, assume that $\Pi = \{\{1, 2\}, \{3\}\}$. Let $e^\Pi = \left(e^{\{1,2\}}, e^{\{3\}}\right)$ and $\phi^\Pi = \left(\phi^{\{1,2\}}, \phi^{\{3\}}\right)$ be the efforts and the transfer scheme that support Π. If, for both players $i = 1, 2$, $U_i\left(\{1, 2\}, \phi^{\{1,2\}}, e^{\{1,2\}}\right) > U_i(\{i\}, e^*)$, then $\{1, 3\}$ can block $\left(\Pi, \phi^\Pi, e^\Pi\right)$ with $\left(\phi^{\{1,3\}}, e^{\{1,3\}}\right)$ where $\phi_1^{\{1,3\}} = \phi_1^{\{1,2\}}$, $\phi_3^{\{1,3\}} = \phi_2^{\{1,2\}}$, $e_1^{\{1,3\}} = e_1^{\{1,2\}}$ and $e_3^{\{1,3\}}$ is slightly higher than $e_2^{\{1,2\}}$. Therefore, $U_i\left(\{1, 2\}, \phi^{\{1,2\}}, e^{\{1,2\}}\right) = U_i(\{i\}, e^*)$ for at least one of $i = 1, 2$. Without loss, assume $U_1\left(\{1, 2\}, \phi^{\{1,2\}}, e^{\{1,2\}}\right) = U_1(\{i\}, e^*)$. Then $\{1, 3\}$ will block Π, by each supplying e^* with 0 transfers because f is strictly increasing.

2. Suppose a stable partition Π exists. Thanks to proposition 1, $\Pi \neq \{\{1\}, \{2\}, \{3\}\}$. Hence without loss of generality, assume that $\Pi = \{\{1, 2\}, \{3\}\}$. Let $e^\Pi = \left(e^{\{1,2\}}, e^{\{3\}}\right)$ be the efforts that support Π. If, for both players $i = 1, 2$, $U_i\left(\{1, 2\}, e^{\{1,2\}}\right) > U_i(\{i\}, e^*)$, by offering a slightly higher effort than player 2's to player 1, players 3 and 1 will block $\left(\Pi, 0, e^\Pi\right)$. Therefore, $U_i(\{1, 2\}, e^{\{1,2\}}) = U_i(\{i\}, e^*)$ for at least one $i = 1, 2$. Without loss, assume $U_1\left(\{1, 2\}, e^{\{1,2\}}\right) = U_1(\{i\}, e^*)$. Then $\{1, 3\}$ will block $\left(\Pi, e^\Pi\right)$ by each supplying e^* since f is strictly increasing.

□

In the above example, even though players 2 and 3 are identical, because efforts are verifiable, player 1 accepts player 3's offer as long it is slightly better than that of player 2. This causes instability.

[3]For more information see Foley (1970), Demange (1987), and Moulin (1987).

By slightly modifying example 1, one obtains an example with no stable partition for the $r = 3$ case. Specifically, if we decrease $\rho(3)$ sufficiently, and $\rho(2)$ slightly, then no stable partition exists when $r = 3$.

Finally, let us remark that the players have no incentive to form a bigger coalition if the credit function decreases fast enough. Therefore, in these cases, each player works on her own. Indeed this is the case for private goods.

Proposition 2. *If teams produce private goods, then*

1. $\Pi = \{\{1\}, \cdots, \{n\}\}$ *is the unique stable partition*
2. $\Pi = \{\{1\}, \cdots, \{n\}\}$ *is the unique stable partition under 0 transfer schemes.*

Proof. Proving the first part yields the second one because, for any one player team, the transfer scheme is 0 due to balancedness. First let us show that $\Pi = \{\{1\}, \cdots, \{n\}\}$ is a stable partition. Let $e_i^* = \arg\max_{e_i} U_i(\{i\}, e_i)$. Under partition Π, each player's utility is $f(\lambda_i e_i^*) - c(e_i^*)$. Suppose team S blocks Π. Then, for some ϕ^S and e^S, $U_i(S, \phi^S, e^S) > U_i(\{i\}, e^*)$ for all $i \in S$. Consequently, $\sum_{i\in S} U_i(S, \phi^S, e^S) > \sum_{i\in S} U_i(\{i\}, e^*)$ or equivalently, $f(\sum_{i\in S} \lambda_i e_i^S) - \sum_{i\in S} c(e_i^S) > \sum_{i\in S}(f(\lambda_i e_i^*) - c(e_i^*))$. We will show this inequality is not satisfied, reaching a contradiction. Observe that the function $\sum_{i\in S}(f(\lambda_i e_i) - c(e_i))$ is maximized if $e_i = e_i^*$ for all $i \in S$. Hence, it suffices to show $f(\sum_{i\in S} \lambda_i e_i^S) < \sum_{i\in S} f(\lambda_i e_i^S)$. Let j be the player with $\lambda_j e_j^S = \max_{i\in S}\{\lambda_i e_i^S\}$. Then, by the concavity of f, $f(\sum_{i\in S} \lambda_i e_i^S) < f(\lambda_j e_j^S) + f'(\lambda_j e_j^S)\sum_{i\neq j\in S} \lambda_i e_i^S$. By the concavity of f, $f'(\lambda_j e_j^S) < f'(\lambda_i e_i^S)$ for all $i \in S$. Hence, $f(\sum_{i\in S} \lambda_i e_i^S) < f(\lambda_j e_j^S) + \sum_{i\neq j\in S}(f'(\lambda_i e_i^S)\lambda_i e_i^S)$. Because $f(0) = 0$ and f is concave, $f'(\lambda_i e_i^S)\lambda_i e_i^S < f(\lambda_i e_i^S)$. This yields $f(\sum_{i\in S} \lambda_i e_i^S) < \sum_{i\in S} f(\lambda_i e_i^S)$.

To show $\Pi = \{\{1\}, \cdots, \{n\}\}$ is a unique stable partition, observe that the proof of $\Pi = \{\{1\}, \cdots, \{n\}\}$ being a stable partition implies that any team S with 2 or more players is blocked by one of its members. Therefore, $\Pi = \{\{1\}, \cdots, \{n\}\}$ is the unique stable partition.

□

4 Team Formation in the Presence of Moral Hazard

In this section, we study the team formation problem under non-verifiable efforts, which is the main interest of this paper.

When efforts are not verifiable, the first question we have to answer is what efforts each team can credibly commit. The answer to this question depends on how freely the players can make transfers among themselves. As in the previous section, suppose the players are about to form teams, leading to partition Π_r. Again each team has to fix its transfer scheme as well as its effort profile. However, not all effort profiles are enforceable, as efforts are not verifiable. Hence, if a team decides to supply a certain effort profile, then it must be credible; in the production stage, no member of the team should find it profitable to deviate from the previously decided effort profile. To model this, suppose team S uses a transfer scheme ϕ^S. Then $\Gamma(S,\phi^S) = \langle S, \mathbb{R}_+^{|S|}, (U_i(S,\phi^S,\cdot))_{i\in S}\rangle$ is a normal form game in which the members of S choose their effort. We assume that the players can use mixed strategies; σ^S stands for a mixed strategy effort profile for team S. By abusing the notations, we use e^S for a pure strategy effort profile for team S. Each game results in a Nash equilibrium. Hence, let $NE(S,\phi^S)$ be the set of Nash equilibrium effort profiles (including mixed) for team S. Therefore, the members of any team S can coordinate on one of the Nash equilibrium effort profiles. Let $NE\left(\Pi_r,\phi^{\Pi_r}\right) = \prod_{S\in\Pi_r} NE(S,\phi^S)$ be the Nash equilibrium effort set for partition Π_r. For the 0 transfer schemes case, we simplify the notations by using $\Gamma(S)$ for $\Gamma(S,0)$, $NE(S)$ for $NE(S,0)$ and $NE\left(\Pi_r\right)$ for $NE\left(\Pi_r,0\right)$.

In order to define the stable partitions we must incorporate the credibility criteria into the definitions of stable partition. As for the verifiable effort case, we define two notions of stable partition.

Definition 2. *In the presence of moral hazard,*

1. *a partition Π_r is stable if, for some ϕ^{Π_r} and $\sigma^{\Pi_r} \in NE(\Pi_r,\phi^{\Pi_r})$, no team S blocks $(\Pi_r,\phi^{\Pi_r},\sigma^{\Pi_r})$ with any ϕ^S and $\sigma^S \in NE(S,\phi^S)$.*
2. *a partition Π_r is stable under 0 transfer schemes if, for some $\sigma^{\Pi_r} \in NE(\Pi_r)$, no team S blocks $(\Pi_r, 0, \sigma^{\Pi_r})$ with $\phi^S = 0$ and any $\sigma^S \in NE(S)$.*

This definition states that in order to block the status quo, a blocking team must specify a transfer scheme and the corresponding credible effort profile which improves all the members of the team over the status quo. This implicitly assumes that the players understand that they cannot base their decision to leave the status quo partition on non-Nash equilibrium effort profiles.

Team Formation under Moral Hazard with Transfer Schemes

To study the existence of stable partitions, one must find the utility possibility set for each team. For private goods, this set is already found in the multi-agent moral hazard literature. In order to use these results, we need to transform our problem into a problem in which the teams produce private goods. To do this, consider any non-private, local goods, i.e., $\rho(s) \neq \frac{1}{s}$ for some $2 \leq s \leq r$. For each team S, let the production function be $\bar{f}^S = |S|\rho(|S|)f$, the credit function be $\bar{\rho}(\cdot)$ where $\bar{\rho}(s) = 1/s$ for all $s = 1, \cdots, n$, and the transfer scheme be $\bar{\phi}^S(\cdot)$ where $\bar{\phi}_i^S(x) = \phi_i^S\left(\frac{x}{\rho(|S|)|S|}\right)$. With this relabeling, the teams produce a private good. As a consequence, we can use the results from the multi-agent moral hazard literature in our setting.

For private goods, Hölmström (1982) shows that, for any balanced transfer scheme, the efficient effort profile cannot be supported as an equilibrium. The main argument behind this result is that some agent is always able to shirk without being identified. However, Legros and Matthews (1993) show that the teams can approximate their efficient level of welfare to any degree using balanced transfer schemes. They construct a transfer scheme and the corresponding mixed equilibrium that approximate the efficient level of production. The incentives are provided as follows: one of the agents (agent 1) should be indifferent between supplying 0 effort or her efficient effort when everyone else supplies their efficient effort. Let agent 1 supply 0 effort with probability ϵ. In addition, if the production outcome is lower than the one corresponding to the case in which all but agent 1 supply their efficient effort while agent 1 supplies 0 effort, then every other player pays a large fine to agent 1. To avoid this large fine, everyone supplies their efficient effort. We present the Legros and Matthews (1993) result in the following lemma.

Lemma 1. *For any $\epsilon > 0$, there exists a balanced transfer scheme ϕ^S such that $\Gamma(S, \phi^S)$ has a mixed equilibrium σ^S according to which expected welfare is within ϵ of the efficient level of welfare:*

$$E_{\sigma^S}(W^S(\tilde{e}^S)) > \bar{W}^S - \epsilon$$

Proof. See Legros and Matthews (1993). □

Lemma 1 has a striking implication for stability: if a stable partition exists, then each player must work alone. The reason is that any team of two or more players cannot agree on any transfer scheme and effort profile, as there is always some other transfer scheme and effort profile that improve everyone over the original one.

Proposition 3. *If a partition Π_r is stable under non-verifiable efforts, then $\Pi_r = \{\{1\}, \cdots, \{n\}\}$. In addition, a stable partition under non-verifiable efforts exists if and only if $\Pi_r = \{\{1\}, \cdots, \{n\}\}$ is a stable partition under verifiable efforts.*

Proof. The first part directly follows from the definition of partition Π_r and lemma 1. To prove the second part, observe that the effort an one person team exerts does not depend on the verifiability of efforts. Hence, the utility one obtains working alone is not affected by the non-verifiability of efforts. Clearly, if no team can block $\Pi_r = \{\{1\}, \cdots, \{n\}\}$ with any transfer scheme and efforts when efforts are verifiable, then no team can block Π_r with any balanced transfer scheme and credible efforts when efforts are not verifiable. On the other hand, if Π_r is blocked under verifiable efforts, then it is blocked under non-verifiable efforts thanks to lemma 1. $\square$

Remarks:

- Proposition 3 implies that moral hazard undermines stability as it significantly reduces the set of stable partitions. For example, for public goods, the set of stable partitions is always empty when efforts are not verifiable. In addition, this negative conclusion is not the artifact of our assumption that the efforts of different players are perfect substitutes. In fact, as long as efforts are bounded below and the production function is differentiable, lemma 1 is satisfied (Legros and Matthews (1993)). Hence, in all these cases, moral hazard affects stability negatively.

- Let us comment on the core under non-verifiable efforts as it is a fundamental concept in cooperative game theory. In our setting, the core under non-verifiable efforts is always empty. The main reason is that the utility possibility set for N is open when efforts are non-verifiable and the team members make only balanced transfers. Hence, there is no hope for the non-emptiness of the core regardless of the kind of goods teams produce.

Coalition Formation under Moral Hazard without Transfer Schemes

As in the previous case, let us start by exploring the utility possibility set for each team. Because the team members cannot make any transfers, the utility possibility set for any team will shrink significantly. In fact, in our setting, the utility possibility set reduces to a single point. We show this in the next lemma.

Lemma 2. *For any game $\Gamma(S)$ there exists a unique Nash equilibrium.*

Proof. First let us show the existence of the Nash equilibria. Consider function $g : \mathbb{R}_+ \to \mathbb{R}_+$ such that $g(e) = \sum_{i \in S} \lambda_i e_i(e)$ where $e_i(e)$ is the solution to $c'(e_i) = \rho(|S|) \lambda_i f'(e)$ for all $i \in S$. Observe that any Nash equilibrium effort vector $\bar{e}^S$ satisfies the following condition: $c'\left(\bar{e}_i^S\right) = \rho(|S|) \lambda_i f'\left(\sum_{j \in S} \lambda_j \bar{e}_j^S\right)$ for all $i \in S$. Hence, if we show that $g(\cdot)$ has a fixed point, then the existence of Nash equilibria is guaranteed. Clearly, $g(\cdot)$ is continuous. In addition, the concavity of $f(\cdot)$, the convexity of $c(\cdot)$ and $c'(0) = 0$ imply that $e_i(0) \geq 0$ for all $i \in S$. Hence, $g(0) \geq 0$.

Furthermore, the concavity of $f(\cdot)$, the convexity of $c(\cdot)$ and the existence of $\bar{e}$ with $\lambda_i f'(\lambda_i \bar{e}) < c'(\bar{e})$ for all i imply that there exists $\tilde{e} > 0$ such that $\tilde{e} > g(\tilde{e})$. Hence, by the intermediate value theorem, there must exist e^* such that $g(e^*) = e^*$. This completes the proof that a Nash equilibrium exists.

To show the uniqueness of the equilibria for each game $\Gamma(S)$, suppose there are two distinct equilibrium efforts $e^S, \bar{e}^S \in NE(S)$. By definition, $e_i^S = \arg\max_{e_i} U_i\left(S, (e_i, e_{-i}^S)\right)$ and $\bar{e}_i^S = \arg\max_{e_i} U_i\left(S, (e_i, \bar{e}_{-i}^S)\right)$ for all $i \in S$. By the first order conditions, $\rho\left(|S|\right)\lambda_i f'\left(\sum_{j\in S}\lambda_j e_j^S\right) = c'\left(e_i^S\right)$ and $\rho\left(|S|\right)\lambda_i f'\left(\sum_{j\in S}\lambda_j \bar{e}_j^S\right) = c'\left(\bar{e}_i^S\right)$ for all $i \in S$. If $\sum_{i\in S}\lambda_i e_i^S = \sum_{i\in S}\lambda_i \bar{e}_i^S$, then the strict convexity of $c(\cdot)$ and the first order conditions imply that $e_i^S = \bar{e}_i^S$ for all $i \in S$. Hence, $e^S = \bar{e}^S$. If $\sum_{i\in S}\lambda_i e_i^S < \sum_{i\in S}\lambda_i \bar{e}_i^S$, then, by the concavity of $f(\cdot)$, $f'(\sum_{i\in S}\lambda_i e_i^S) \geq f'\left(\sum_{i\in S}\lambda_i \bar{e}_i^S\right)$. Then the convexity of $c(\cdot)$ and the first order conditions imply that $e_i^S \geq \bar{e}_i^S$ for all $i \in S$. This contradicts $\sum_{i\in S}\lambda_i e_i^S < \sum_{i\in S}\lambda_i \bar{e}_i^S$. □

Thanks to the uniqueness of the Nash equilibria, the players know their utility from each team. Hence, the players have a well-defined preference relation over the teams which include her. Consequently, when efforts are non-verifiable, our team formation problem under 0 transfer schemes is a hedonic game defined by Bogomolnaia and Jackson (2002).

To study the existence of stable partitions, it would be ideal if one knows how a player ranks any two teams by the utility the teams bring to her. Unfortunately, this is impossible without specifying the credit and production function. However, we can characterize a player's ranking of two teams with equal size in some cases. Specifically, the following lemma shows that each player prefers to replace a less productive member (not herself) of her team with a more productive player. The reason is that a more productive player supplies more (productivity weighted) effort at equilibrium than a less productive one does. Before presenting the lemma, let us fix some notations: $U_i(S)$, and $e_\lambda(S)$ denote the equilibrium utility of player $i \in S$, and equilibrium weighted effort, respectively.

Lemma 3. *Consider any S with $|S| \geq 2$ and $\bar{S} = S \backslash \{i\} \cup \{\bar{i}\}$ where $i \in S$. Then, for all $j \neq i$ in S, $U_j(\bar{S}) \gtreqqless U_j(S)$ if and only if $\lambda_{\bar{i}} \gtreqqless \lambda_i$.*

Proof. Let e^S and $e^{\bar{S}}$ be the Nash equilibrium effort profiles for teams S and $\bar{S}$, respectively. First let us show that $e_\lambda(\bar{S}) \gtreqqless e_\lambda(S)$ if and only if $\lambda_{\bar{i}} \gtreqqless \lambda_i$. If $\lambda_{\bar{i}} = \lambda_i$, then the uniqueness of Nash equilibria yields that $e_\lambda(\bar{S}) = e_\lambda(S)$. If $e_\lambda(\bar{S}) \leq e_\lambda(S)$ when $\lambda_{\bar{i}} > \lambda_i$, then by the concavity of $f(\cdot)$, $f'\left(e_\lambda(\bar{S})\right) \geq f'\left(e_\lambda(S)\right)$. Then the convexity of $c(\cdot)$ and the first order conditions imply that $e_j^{\bar{S}} \geq e_j^S$ for all $j \in S\backslash\{i\}$ and $e_{\bar{i}}^{\bar{S}} \geq e_i^S$. Therefore, $\sum_{j\in\bar{S}}\lambda_j e_j^{\bar{S}} > \sum_{j\in S}\lambda_j e_j^S$, contradicting $e_\lambda(\bar{S}) \leq e_\lambda(S)$. This shows that if $\lambda_{\bar{i}} > \lambda_i$, then $e_\lambda(\bar{S}) > e_\lambda(S)$. Similarly, one can show that if $\lambda_{\bar{i}} < \lambda_i$, then $e_\lambda(\bar{S}) < e_\lambda(S)$. Hence, $e_\lambda(\bar{S}) \gtreqqless e_\lambda(S)$ if and only if $\lambda_{\bar{i}} \gtreqqless \lambda_i$.

Now we show that $e_\lambda(\bar{S}) \gtreqqless e_\lambda(S)$ if and only if $U_j(\bar{S}) \gtreqqless U_j(S)$. If $e_\lambda(\bar{S}) = e_\lambda(S)$, then the strict convexity of $c(\cdot)$ yields $e_j^{\bar{S}} = e_j^S$ for all $j \in S\backslash\{i\}$. Therefore, $U_j(\bar{S}) = U_j(S)$ for all $j \in S\backslash\{i\}$. If $e_\lambda(\bar{S}) > e_\lambda(S)$, then the strict convexity of $c(\cdot)$ yields $e_j^{\bar{S}} \leq e_j^S$ for all $j \in S\backslash\{i\}$. Since f is strictly increasing, $U_j(\bar{S}) > U_j(S)$ for all $j \in S\backslash\{i\}$. Similarly, one can show that if $e_\lambda(\bar{S}) < e_\lambda(S)$, then $U_j(\bar{S}) < U_j(S)$ for all $j \in S\backslash\{i\}$. □

Next lemma states that by recruiting a player to one's current team, the player's equilibrium effort decreases. Hence, the players always save their cost (from efforts) by expanding their current team.

Lemma 4. *If $\bar{S} = S \cup \{\bar{i}\}$, then $e_i^{\bar{S}} \leq e_i^S$ where e_i^S and $e_i^{\bar{S}}$ are player i's equilibrium efforts corresponding to team S and $\bar{S}$.*

Proof. Let $e_\lambda(T, \rho)$ and $e^T(\rho)$ denote the total weighted effort and the equilibrium effort profile for team T when $\rho(|T|) = \rho$. Slightly modifying the proof of lemma 3, one can prove that $e_\lambda(\bar{S}, \rho(|S|)) > e_\lambda(S, \rho(|S|))$. Then the convexity of $c(\cdot)$ and the first order conditions imply that $e_i^{\bar{S}}(\rho(|S|)) \leq e_i^S(\rho(|S|))$ for all $i \in S$. Now let us show that if $\rho(|\bar{S}|) \leq \rho(|S|)$ then $\rho(|\bar{S}|)f'\left(e_\lambda(\bar{S}, \rho(|\bar{S}|))\right) \leq \rho(|S|)f'\left(e_\lambda(S, \rho(|S|))\right)$. If this is not true, then the first order conditions and the convexity of $c(\cdot)$ imply that $e_i^{\bar{S}}(\rho(|\bar{S}|)) \geq e_i^{\bar{S}}(\rho(|S|))$. Then $e_\lambda(\bar{S}, \rho(|\bar{S}|)) \geq e_\lambda(\bar{S}, \rho(|\bar{S}|))$. Then the concavity of $f(\cdot)$ implies that $\rho(|\bar{S}|)f'\left(e_\lambda(\bar{S}, \rho(|\bar{S}|))\right) \leq \rho(|S|)f'\left(e_\lambda(S, \rho(|S|))\right)$ which is a contradiction.

If $\rho(|\bar{S}|)f'\left(e_\lambda(\bar{S}, \rho(|\bar{S}|))\right) \leq \rho(|S|)f'\left(e_\lambda(S, \rho(|S|))\right)$, then $e_i^{\bar{S}}(\rho(|\bar{S}|)) \leq e_i^S(\rho(|S|))$ thanks to the first order conditions and the convexity of $c(\cdot)$. □

Now we are ready to present our results on the existence of the stable partition and its characterization. Without loss of generality, we assume that the players are indexed by productivity in decreasing order, i.e., if $i < j$, then $\lambda_i \geq \lambda_j$.

Our first result states that if teams produce public goods, then a stable partition exists. The key reason behind this result is the lack of punishment for forming a bigger team due to the public nature of the good. Therefore, a player can only benefit by recruiting an additional player to her current team since the total weighted effort increases while each player's equilibrium effort decreases. Therefore, every team is of size r except maybe one. To be specific, the most productive r players are matched, and then the next most productive r players are matched and so on. This also implies that if $r \geq n$, then the unique stable partition is $\Pi_r = \{N\}$.

Proposition 4. *If teams produce public goods, then there exists a stable partition. Furthermore, if $\lambda_i \neq \lambda_j$ for all $i \neq j$, then there is a unique stable partition. Moreover, every stable partition is assortative.*

Proof. To prove the existence consider the following canonical partition Π_r resulting from the following algorithm.

1. Player 1 picks a team respecting the quota. If there are multiple best matches then she picks the team whose total indices are the smallest (There will not be any case in which one player's best matches have the same total indices).

2. The lowest indexed *unmatched* player picks a team out of the remaining players respecting the quota.

3. Continue with step 2 until all the players are matched.

Claim. Partition Π_r is stable.

Proof of the Claim. Consider the above canonical algorithm. Slightly modifying the proof of lemma 3, one obtains that if $\rho(\cdot)$ is constant, then, for each $i \in S$, $U_i(S \cup \{j\}) > U_i(S)$. Combining this with lemma 3, one obtains that player 1 chooses to match with the most productive r players (including herself) if $r \leq n$. Observe here that in step 2 of the algorithm the player who chooses a team is the most productive player out of the unmatched players. As player 1, she will choose to match with the most productive unmatched r players if $r \leq n/2$. This logic extends to all the remaining teams. Let $\{S_1, ..., S_J\}$ be the teams resulting from the canonical algorithm in which they are indexed by the order they were formed.

To prove the stability, observe that everyone in team 1 obtains her best possible match, so no one would be a part of any blocking team. Given this, anyone in the second team would not be a part of any blocking team. Continuing with this logic, we can see that no team blocks partition Π_r.

When the productivities of no two players are the same, S_1 is the unique best match for anyone in S_1. Given this, S_2 is the unique best match out of $N \backslash S_1$ for anyone in S_2 and so on. Therefore, Π_r is obviously the unique stable partition.

Now let us prove that the stable partitions are assortative. Suppose that $\mathcal{B}_r$ is a stable partition. Let $\mathcal{B}_r = \{T_1, ..., T_K\}$ where the teams are indexed by the total productivity coefficients in decreasing order. Observe here that there must be a bijection $\pi_1 : S_1 \to T_1$ such that $\lambda_i = \lambda_{\pi_1(i)}$ for all $i \in S_1$; otherwise, S_1 must block $\mathcal{B}_r$ because S_1 is the best match for all of its members. Similarly, there exists $\pi_j : S_j \to T_j$ such that $\lambda_i = \lambda_{\pi_j(i)}$ for all $i \in S_j$. This implies that $\mathcal{B}_r$ is assortative. $\square$

Remark:

- Our team formation game for public goods satisfies the *top team property* defined by Banerjee et al. (2001). This property says that for any subset $V \subset N$ there exists team $S \subset V$ such that each member of S (weakly) prefers S to any team $T \subset V$. To see that our problem satisfies the top team property, observe that team S, which consists of r players with the highest productivity out of team V, brings no less utility to its members than any other team $T \subset V$.

Now we consider the case in which credit function ρ is not constant. In this case, recruiting an additional player to an already formed team has three effects on each existing member: (1) each player earns less credit (2) each player exerts less effort, hence saves some cost of efforts (lemma 4) and (3) the equilibrium weighted effort changes, affecting each player's consumption of local goods. The direction of the equilibrium weighted effort change depends on the credit function and the cost function. However, one can imagine that if ρ is decreasing fast enough, then the first effect must dominate the other two, eliminating the incentives to form a bigger team. Indeed, this is the case for private goods we already know from proposition 2. Thus, for the private goods case, no player finds it profitable to work with other players who have a lower productivity. Therefore, every player works by herself which is a stark contrast to the public good case in which the players form teams of a size as big as possible.

Now we will consider the case in which the local goods are neither public nor private. In this case, if the quota does not exceed 2, then one can prove the existence of stable partitions. The reason behind this result is that the most productive player obtains the highest utility by working alone. However, if two players work together, then the least productive player obtains a better utility than the other. Therefore, if the most productive player offers someone to partner, then the other must accept the offer. Exploiting this property, we design a canonical algorithm in which the most productive player chooses her match and then the most productive unmatched player chooses her match from the unmatched players and so on. Now let us present the result.

Proposition 5. *A stable partition Π_r exists if $r = 2$. In addition, stable partitions are assortative.*

Proof. To prove the existence of stable partitions, consider partition Π_r which is the result of the algorithm used in the proof of proposition 4.

Claim. Π_r is stable.

Proof of the claim. It suffices to prove that no player in $\Pi_r(1)$ is part of a blocking team. If $\Pi_r(1) = \{1\}$, then, by construction, $U_1(\{1\}) \geq U_1(\{1, 2\})$. Therefore, player 1 is not a part of any blocking team.

If $|\Pi_r(1)| = 2$, then due to lemma 3 and the construction of the canonical mechanism, $\Pi_r(1) = \{1,2\}$. By lemma 3, no team of size 2 can block $\{1,2\}$. By construction, $U_1(\{1\}) < U_1(\{1,2\})$, hence, player 1 will not block $\{1,2\}$. Now, to show player 2 will not block $\{1,2\}$ by herself, it suffices to show that (1) $U_2(\{1,2\}) \geq U_1(\{1,2\})$ and (2) $U_2(\{2\}) \leq U_1(\{1\})$. To see (1), observe that the first order conditions yield $\rho\,(2)\,\lambda_i f'\left(\lambda_1 e_1 + \lambda_2 e_2\right) = c'\left(e_i\right)$ for $i = 1, 2$. Therefore, $\frac{c'(e_1)}{\lambda_1} = \frac{c'(e_2)}{\lambda_2}$. Because c is strictly convex and $\lambda_1 \geq \lambda_2$, $e_1 \geq e_2$. As a result,

$$U_2(\{1,2\}) = \rho\,(2)\,f\left(\lambda_1 e_1 + \lambda_2 e_2\right) - c\left(e_2\right) \geq \rho\,(2)\,f\left(\lambda_1 e_1 + \lambda_2 e_2\right) - c\left(e_1\right) = U_1(\{1,2\})$$

To see (2), observe that $\max_{e_1}\left\{f(\lambda_1 e_1) - c\left(e_1\right)\right\} \geq \max_{e_2}\left\{f(\lambda_2 e_2) - c\left(e_2\right)\right\}$ because, for all e, $f(\lambda_1 e) - c\left(e\right) \geq f(\lambda_2 e) - c\left(e\right)$ as $\lambda_1 \geq \lambda_2$.

Now let us show that all stable partitions are assortative when $r = 2$.
Consider any non-assortative partition $\mathcal{B}$. Since $\mathcal{B}$ is non-assortative, there must exist $S, T \in \mathcal{B}$ such that $\max_{i \in S} \lambda_i > \min_{i \in T} \lambda_i$ and $\max_{i \in S} \lambda_i > \min_{i \in T} \lambda_i$. Let $\bar{i}/\underline{i}$ and $\bar{j}/\underline{j}$ be the lowest/highest indexed players in team S and T, respectively. Without loss of generality, we assume $\lambda_{\bar{i}} \geq \lambda_{\bar{j}}$. Consequently, $|S| = 2$; otherwise $\min_{i \in S} \lambda_i \geq \max_{i \in T} \lambda_i$. Now we show that $\mathcal{B}$ is blocked by some team.
Let $U_{\bar{i}}(\{\bar{i}\}) \geq U_{\bar{i}}(\{\bar{i}, \bar{j}\})$. Then $\{\bar{i}\}$ blocks $\mathcal{B}$ because $U_{\bar{i}}(\{\bar{i}\}) \geq U_{\bar{i}}(\{\bar{i}, \bar{j}\}) > U_{\bar{i}}(S)$ due to lemma 3 as $\lambda_{\bar{j}} > \lambda_{\underline{i}}$.
Let $U_{\bar{i}}(\{\bar{i}\}) < U_{\bar{i}}(\{\bar{i}, \bar{j}\})$. Then by lemma 3, $U_{\bar{i}}(\{\bar{i}, \bar{j}\}) > U_{\bar{i}}(S)$ as $\lambda_{\bar{j}} > \lambda_{\underline{i}}$. If $|T| = 2$, by lemma 3, $U_{\bar{j}}(\{\bar{i}, \bar{j}\}) > U_{\bar{j}}(T)$ as $\lambda_{\bar{i}} > \lambda_{\underline{j}}$. As a result, $\{\bar{i}, \bar{j}\}$ blocks $\mathcal{B}$. Lastly, if $|T| = 1$, then, from the first part of this proof, $U_{\bar{j}}(\{\bar{j}\}) < U_{\bar{j}}(\{\bar{i}, \bar{j}\})$ as $\lambda_{\bar{i}} > \lambda_{\bar{j}}$ and $U_{\bar{i}}(\{\bar{i}\}) < U_{\bar{i}}(\{\bar{i}, \bar{j}\})$. Hence, $\{\bar{i}, \bar{j}\}$ blocks $\mathcal{B}$, completing the proof. $\square$

Let us remark some interesting features of the stable partitions with quota $r = 2$.

1. Our team formation game with quota $r = 2$ satisfies the top team property. Therefore, to prove the existence of stable partitions, we could have shown first that our team formation game satisfies the top team property and then obtained the existence as a corollary of theorem 1 in Banerjee et al. (2001). However, we chose the current approach because the canonical algorithm used in the proof has certain value as it demonstrates how to obtain a stable partition.

2. Our team formation game with quota $r = 2$ satisfies the no odd rings condition defined by Chung (2000).

3. The size of teams resulting from the algorithm could differ from team to team. The following example illustrates this point.

   Example 2. *Let the set of players be* $\{1, 2, 3, 4, 5\}$ *and the probability of a project succeeding be* $f(e) = \frac{\exp(e)}{1+\exp(e)}$. *Suppose* $c(e_i) = \frac{1}{2}e_i^2$, $\lambda_1 = \sqrt{70}$, $\lambda_2 = \lambda_3 = \sqrt{50}$, $\lambda_4 = \sqrt{20}$, *and* $\lambda_5 = \sqrt{10}$. *Finally, let* $\rho(1) = 1$, $\rho(2) = 0.95$ *and* $r = 2$.
   In this game, the stable partition is $\{\{1, 2\}, \{3\}, \{4, 5\}\}$. *Even though player* 3 *could match with player* 4, *she decides to stay unmatched. When matched with player* 4, *there are two contradicting effects for player* 3. *The first one is the utility gain resulting from player 4's positive effort and the second one is the utility loss because of a decreased* ρ. *When the difference between productivity coefficients is significant, player* 4 *puts too low of an effort relative to that of player 3 when matched with player* 3. *Therefore, the utility gain from player 4's positive effort is not enough to overcome the utility loss from decreased* ρ. *However, for player* 4 *partnering with player* 5 *is beneficial since the effort player* 5 *supplies is comparable to that of player* 4.

The existence of stable partitions when $r > 2$ depends on the cost function. For example, lemma 6, which we will see shortly, proves that the existence of stable partitions is guaranteed when $r \leq 4$ if the cost function is quadratic. Let us take a moment to explore the quadratic cost case which provides an additional structure that we will exploit in constructing an example with no stable partition. An interesting feature of the quadratic cost function is that the equilibrium weighted effort for team S depends only on $\sum_{i \in S} \frac{\lambda_i^2}{a}$. This allows us to derive a player's preference ranking of the teams with same size. Before presenting the result, let us fix some notations; $t_i = \frac{\lambda_i^2}{a}$, $t(S) = \sum_{j \in S} t_j$, and $t_{-i}(S) = \sum_{j \neq i, j \in S} t_j$.

Lemma 5. *If* $c(e) = \frac{a}{2}e^2$, *then*

a. *for any game* $\Gamma(S)$, $e_\lambda(S)$ *satisfies the following condition.*

$$e_\lambda(S) = \rho\left(|S|\right) t(S) f'\left(e_\lambda(S)\right) \tag{2}$$

Furthermore, player i*'s utility function at the equilibrium is*

$$U_i(S) = \rho\left(|S|\right) f\left(e_\lambda(S)\right) - \frac{t_i}{2}\left(\frac{e_\lambda(S)}{t(S)}\right)^2 \tag{3}$$

b. *Suppose teams* S *and* T *have the same size* ($|S| = |T|$) *and* $S \cap T \neq \emptyset$. *Then for player* $i \in S \cap T$, $U_i(S) \leq U_i(T)$ *if and only if* $t_{-i}(S) \leq t_{-i}(T)$.

Proof. See Appendix. □

Proposition 6 proves the existence of stable partitions for the $r \leq 4$, quadratic cost case. The reason for the existence of stable partition is that the most productive player in any team is not expendable when the quota is low. For example, suppose players 1, 2 and 3 team up and let u_1, u_2 and u_3 be the corresponding utilities for players 1, 2, and 3. The only potential conflict between players 1 and 2 is that player 1 might prefer to recruit player 4 while player 2 does not. Then, player 2 would consider replacing player 1 with player 4 from the original team which is costly to player 2 because player 4 is less productive than player 1. This replacement cost for player 2 exceeds the cost player 2 incurs if player 4 joins team $\{1, 2, 3\}$. By exploiting this property, we consider the canonical algorithm in which the most productive player chooses her team and then the most productive unmatched player chooses her team and so on.

Proposition 6. *If* $r \leq 4$ *and* $c(e) = \frac{a}{2}e^2$*, then a stable partition exists. In addition, the stable partitions are assortative.*

Proof. See Appendix. □

Remark:
Our team formation game with quota $r \leq 4$ and quadratic cost function may not satisfy the top team property but *does* satisfy the *weak* top team property defined in Banerjee et al. (2001).

To identify the effects of moral hazard on stability, we have investigated the existence of stable partitions under 0 transfer schemes. Now let us summarize all the results of this section in the following theorem.

Theorem 1. *In the presence of moral hazard, there exists a stable partition under 0 transfer schemes if*

1. *teams produce public or private goods*
2. $r = 2$
3. *the cost function is quadratic and* $r \leq 4$

In all above cases, all stable partitions are assortative.

Based on theorem 1 and example 1, we conclude that the combination of moral hazard and the lack of transfer schemes greatly increases the set of stable partitions in certain cases. For example, for public goods, a stable partition always exists regardless of the quota. For non-public goods, the quota plays an important role for whether stable partition exists or not: if $r = 2$ or if $r \leq 4$ and the cost function is quadratic, there is always a stable partition. For all these cases, example 1 and proposition 3 demonstrate that we need both moral hazard and the lack of transfer schemes to guarantee the existence of stable partitions. Therefore, we conclude that the combination of moral hazard and the lack of transfer schemes positively affects stability in a large class of games. But this gain in stability comes with a considerable loss in efficiency because the combination of moral hazard and the lack of transfer scheme deteriorates efficiency the most.

Now let us discuss the reason why in some cases there is no stable partition when efforts are verifiable but there is one when efforts are not verifiable and there is no transfer scheme. The biggest difference between these cases is the expectations of the other players' efforts after a team is formed. Suppose S is formed. To break down team S, some players not in S must recruit some members of S to form a new team. To accomplish this, they must propose an effort profile which improves the members of the new team. For the verifiable case, this task is relatively easy as all efforts are enforceable. However, for the non-verifiable case, everyone expects the new team to exert only credible effort profile in the production stage. In our setting, there is only one credible effort profile, and if this does not cause instability then team S will form.

Another interesting feature of our model is the assortativeness property of stable partitions in the presence of moral hazard. This seems to be observed in life. For example, the top tier researchers coauthor with one another while the more modest researchers work with each other.

We now show that the combination of moral hazard and lack of transfer schemes does not necessarily guarantee the existence of stable partitions. We demonstrate this point in the following example.

Example 3. *Let the set of players be* $\{1, 2, \cdots, 500\}$*,* $f(e) = 1 - \frac{1}{(1+e)^3}$ *and* $c(e_i) = \frac{1}{2}e_i^2$*. In addition,* $t_1 = 101$*,* $t_2 = t_3 = t_4 = 100$ *and* $t_5 = \cdots = t_{500} = 99$*. Finally, let* $\rho(1) = \rho(2) = \rho(3) = 1$ *but* $\rho(j)$ *where* $j \geq 4$ *satisfies the following condition* $U_2(\{1, 2, 3, 4\}) = \cdots = U_2(\{1, 2, 3, 4, \cdots, k\}) = U_2(\{2, 3, 4\}) - \epsilon$ *where* $k = 5, \cdots, 500$ *and* ϵ *is an arbitrarily small, positive number. If there is no quota restriction, then no stable partition exists when* $\epsilon \to 0$*.*

Proof. [4] Contrary to the claim, suppose there exists a stable partition Π. Clearly, $\Pi(1) \geq 3$ as $\rho(1) = \rho(2) = \rho(3) = 1$. There are several cases to consider:

1. $|\Pi(1)| \geq 4$ and at least one of players 2, 3 or 4 is in $\Pi(1)$
2. $|\Pi(1)| \geq 4$ and none of players 2, 3, or 4 is in $\Pi(1)$

[4] Due to the lack of closed form solutions, some of the calculations are performed in Matlab.

3. $|\Pi(1)| = 3$ and at least 2 of players 2, 3, and 4 are not in $\Pi(1)$

4. $|\Pi(1)| = 3$ and 2 of players 2, 3, and 4 are in $\Pi(1)$

1. Pick any $i = 2, 3, 4$, and let $\Pi(i) = S$. If $|S| \leq 3$, then $1 \notin S$ because $|\Pi(1)| \geq 4$ in this case. Consequently, $t_2 + t_3 + t_4 > t^S$, implying $U_i(\{2, 3, 4\}) > U_i(S)$ because $\rho(1) = \rho(2) = \rho(3) = 1$. If $|S| \geq 4$, then consider $S_1 = \{1, 2, 3, 4, \cdots, |S|\}$. By construction, $U_i(\{2, 3, 4\}) > U_i(S_1)$. In addition, $U_i(S_1) \geq U_i(S)$ thanks to lemma 5.b. Therefore, player i strictly prefers $\{2, 3, 4\}$ to S. Hence, Π is blocked by $\{2, 3, 4\}$

2. When $\epsilon \to 0$, $U_1(\{1, 2, 3\}) = 0.980299045$ and $U_1(\{1, 5, \cdots, 500\}) = 0.980299041$. In addition, $U_1(\{1, 5, 6, \cdots, k\})$ increases with $k \geq 6$ from a matlab calculation. Therefore, $U_1(\Pi(1)) < U_1(\{1, 2, 3\})$. At the same time, $U_j(\{1, 2, 3\}) > U_j(\{2, 3, 4\})$ for any $j = 2, 3$ thanks to lemma 5.b. In addition, $U_j(\{2, 3, 4\}) \geq U_j(S)$ for any $S \neq \{1, 2, 3\})$ and $S \ni j$ as we noted earlier. Therefore, $U_j(\Pi(j)) < U_j(\{1, 2, 3\})$ where $j = 2, 3$. Hence, Π is blocked by $\{1, 2, 3\}$.

3. In this case, $U_1(\Pi(1)) < U_1(\{1, 2, 3\})$ thanks to lemma 5.b. At the same time, as noted earlier, $\{1, 2, 3\}$ brings a better payoff to players 2 and 3 than any other team. Hence, Π is blocked by $\{1, 2, 3\}$.

4. Without loss of generality, assume $\Pi(1) = \{1, 2, 3\}$. As $\epsilon \to 0$, $U_1(\{1, 4, 5, \cdots, 500\}) = 0.98029905$. Hence, $U_1(\{1, 4, \cdots, 500\}) > U_1(\{1, 2, 3\})$. In addition, a matlab calculation yields $U_4(\{1, 4, \cdots, 500\}) = 0.98029907 > U_4(S)$ for any team $S \ni 4$ but $1 \notin S$, and $U_j(\{1, 4, \cdots, 500\}) = 0.980299091 > U_j(T)$ where $j = 5, \cdots, 500$, $T \ni j$ and $1 \notin T$. Hence, Π is blocked by team $\{1, 4, \cdots, 500\}$.

□

The key difference between the $r \leq 4$ and the $r > 4$ cases is that the most productive player wants to form such a large team that is blocked by a smaller team. Hence, the partition resulting from the algorithm used in proposition 4 is blocked. To see the intuition, suppose players 1, 2 and 3 ($t_1 > t_2 > t_3$) form a team and let u_1, u_2 and u_3 be the corresponding utilities. The only potential conflict between players 1 and 2 is that player 1 might prefer to bring other players while player 2 does not. In addition, the larger the team size is, the smaller the utility difference between players 2 and 1 will be. Therefore, if player 1 wants to form a very large team which yields a slightly greater utility than u_1 to player 1, then player 2's utility will be in a close neighborhood of u_1. But u_1 is lower than u_2. Therefore, if player 2's utility from team $\{2,3,4\}$ is strictly greater than u_1, then $\{2,3,4\}$ blocks the large team player 1 prefers. Example 3 is constructed based on this intuition.

Based on theorem 1, one might expect that the stable partitions are always assortative. However, this turns out not to be true as the following example illustrates.

Example 4. *Let the set of players be* $\{1,2,\cdots,550\}$*,* $f(e) = 1 - \frac{1}{(1+e)^3}$ *and* $c(e_i) = \frac{a}{2}e_i^2$*. In addition,* $t_1 = 101$*,* $t_2 = t_3 = t_4 = 100$ *and* $t_5 = \cdots = t_{500} = 99$*. Finally, let* $\rho(1) = \rho(2) = \rho(3) = 1$ *but* $\rho(j)$ *where* $j \geq 4$ *satisfies the following condition* $U_2(\{1,2,3,4\}) = \cdots = U_2(\{1,2,3,4,\cdots,k\}) = U_2(\{2,3,4\}) - \epsilon$ *where* $k = 6,\cdots,550$ *and* ϵ *is an arbitrarily small, positive number. If there is no quota restriction, then* $\Pi = \{\{2,3,4\}\{1,5,\cdots,550\}\}$*.*

Proof. [5] Due to the proof of example 3, we need to make sure that $U_1(\{1,5,\cdots,550\}) > U_1(\{1,2,3\})$ and team $\{1,5,\cdots,550\}\}$ is not blocked by any team. The former is easily obtained as a matlab calculation yields that $U_1(\{1,5,\cdots,550\}) = 0.980299047$ and $U_1(\{1,2,3\}) = 0.980299046$. The latter is true because, according to a matlab calculation, $U_1(\{1,5,6,\cdots,k\})$ and $U_5(\{1,5,6,\cdots,k\})$ increases with $k \geq 7$, $U_5(\{5,6,\cdots,k\}) < U_5(\{1,5,6,\cdots,k-1\})$ where $k \geq 8$ thanks to lemma 5.b, and $U_5(\{5,6,7\}) < U_5(\{1,5,6,7\})$. $\square$

This example shows that the assortativeness property is not generalized. Hence, we conclude that assortative team structure emerges if the good is public or if the quota is 2. In other cases, the stable team structures may not be assortative.

5 Conclusion

In this paper, we have studied the team formation problem under moral hazard. First, we show that moral hazard significantly worsens stability when the team members can make balanced transfers only. This result implies the emptiness of the core. However, if the team members cannot make transfers, then moral hazard positively affects stability in a large class of games. Therefore, even though the lack of transfer scheme and moral hazard hurts the efficiency the most, they are desirable in a large class of games.

[5] Due to the lack of closed form solutions, some of the calculations are done in Matlab.

References

Suryapratim Banerjee, Hedeo Konishi, and Tayfun Sönmez. Core in a simple coalition formation game. *Social Choice Welfare*, 18, 2001.

Anna Bogomolnaia and Matthew O. Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38, 2002.

James M. Buchanan. An economic theory of clubs. *Economics*, 32(125):1–14, 1965.

Kim-Sau Chung. On the existence of stable roommate matchings. *Games and Economic Behavior*, 33(2):206–230, November 2000.

Kaniska Dam and David Prez-Castrillo. The principal-agent matching market. *Advances in Theoritical Economics*, 2(1), 2006.

Gabrielle Demange. Nonmanipulable cores. *Econometrica*, 55(5):1057–74, September 1987.

Duncan K. Foley. Lindahl's solution and the core of an economy with public goods. *Econometrica*, 38(1):66–72, January 1970.

David Gale and Lloyd Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1), 1962.

Bengt Hölmström. Moral hazard in teams. *Bell Journal of Economics*, 13(2):324–340, Autumn 1982.

Patrick Legros and Steven A Matthews. Efficient and nearly-efficient partnerships. *Review of Economic Studies*, 60(3):599–611, 1993.

Herve Moulin. Egalitarian-equivalent cost sharing of a public good. *Econometrica*, 55(4):963–76, July 1987.

Todd Sandler and John T. Tschirhart. The economic theory of clubs: An evaluative survey. *Journal of Economic Literature*, 18(4):1481–1521, 1980.

Herbert E. Scarf. The core of an n person game. *Econometrica*, 35(1):50–69, 1967.

Jimmy J. M. Tan. A necessary and sufficient condition for the existence of a complete stable matching. *Journal of Algorithms*, 12(1):154–178, 1991.

Appendix

In order to prove proposition 1, we will transform our team formation problem into the setting of cooperative game defined in Scarf (1967); then we proceed to show this game is balanced in the sense of Scarf (1967). Since the core for the balanced game is not empty, the balancedness implies that partition $\Pi_r = \{N\}$ is the only stable partition. Let us start with some definitions.

A cooperative game is a pair $\langle N, (V^S)_{S\subseteq N}\rangle$ where $V^S \subset \mathbb{R}^{|S|}$ is the set of utility vectors for team S. It is assumed that V^S is closed and every v^S is in V^S if $v^S \ll u^S$ for some $u^S \in V^S$.
Let the boundary of utility vector set be $\partial V^S \equiv \{v^S \in V^S : \nexists u^S \in V^S \text{ s.t. } u^S \gg v^S\}$. The projection for team S is $\pi^S : \mathbb{R}^n \to \mathbb{R}^{|S|}$ satisfying $\pi_i^S(y^N) = y_i^N$ for all $i \in S$. The domain for the projection for team S could be the set efforts or the set of utility vectors.

Definition 3. *A utility vector $u \in V^N$ is in the core of game $\langle N, (V^S)_{S\subseteq N}\rangle$ if $\pi^S(u) \notin V^S \backslash \partial V^S$ for any $S \subseteq N$.*

Denote the possible set of teams by $\mathcal{N} = 2^N \backslash \emptyset$. Let $\mathcal{B}$ be a subset of $\mathcal{N}$ and let $\mathcal{B}_i = \{S \in \mathcal{B} : i \in S\}$. Set $\mathcal{B}$ is balanced if there exists non-negative weights w such that $\sum_{S\in\mathcal{B}_i} w(S) = 1$ for all $i \in N$.

Definition 4. *An n person game $\langle N, (V^S)_{S\subseteq N}\rangle$ is balanced if for every balanced set $\mathcal{B}$, a utility vector u is in V^N if $\pi^S(u) \in V^S$ for all $S \in \mathcal{B}$.*

Proof of Proposition 1. **Part 1.** In our team formation problem with transfers, we set $V^S \equiv \{u^S \in \mathbb{R}^{|S|} : \exists e^S \,\&\, \phi^S \text{ s.t. } u_i^S \leq U_i(S, \phi^S, e^S)\}$. It can easily be shown that V^S is closed and $v^S \in V^S$ if $v^S \leq u^S$ for some $u^S \in V^S$, so our team formation problem with transfers is a cooperative game.
Claim. Team formation problem with transfers is a balanced game.
Proof of the Claim. We need to prove that $u \in \mathbb{R}^N$ is V^N if, for any balanced set $\mathcal{B}$, $\pi^S(u) \in V^S$ for all $S \in \mathcal{B}$.

Fix a balanced set $\mathcal{B}$. Let w be a corresponding balancing weight to $\mathcal{B}$. Consider any u such that $\pi^S(u) \in V^S$ for all $S \in \mathcal{B}$. This means that there exist ϕ^S and e^S, such that $U_i(S, \phi^S, e^S) > u_i$ for all $S \in \mathcal{B}$ and for all $i \in S$. Since $\sum_{S\in\mathcal{B}_i} w(S) = 1$ for all $i \in N$, $\sum_{S\in\mathcal{B}_i} w(S) U_i(S, \phi^S, e^S) \geq u_i$.

Consider an effort profile $\bar{e} \in \mathbb{R}^n$ and a transfer scheme $\bar{\phi} : \mathbb{R}_+ \to \mathbb{R}^n$ such that $\bar{e}_i = \sum_{S\in\mathcal{B}_i} w(S) e_i^S$ and $\bar{\phi}_i(\cdot) = \sum_{S\in\mathcal{B}_i} w(S) \phi_i^S \left(f \left(\sum_{j\in S} \lambda_j e_j^S \right) \right)$. Observe that $\bar{\phi}_i$ is a constant function. We will show that $\sum_{i\in N} \bar{\phi}_i \leq 0$ and $U_i(N, \bar{\phi}, \bar{e}) \geq \sum_{S\in\mathcal{B}_i} w(S) U_i(S, \phi^S, e^S)$, proving $u \in V^N$.

To show $\sum_{i\in N}\bar{\phi}_i \le 0$, observe that

$$\sum_{i\in N}\bar{\phi}_i = \sum_{i\in N}\sum_{S\in\mathcal{B}_i} w(S)\phi_i^S\left(f\left(\sum_{j\in S}\lambda_j e_j^S\right)\right) = \sum_{S\in\mathcal{B}} w(S)\sum_{i\in S}\phi_i^S\left(f\left(\sum_{j\in S}\lambda_j e_j^S\right)\right).$$

Since for all S, $\sum_{i\in S}\phi_i^S\left(f\left(\sum_{j\in S}\lambda_j e_j^S\right)\right) \le 0$, it must be $\sum_{i\in N}\bar{\phi}_i \le 0$.

To show $U_i(N,\bar{\phi},\bar{e}) \ge \sum_{S\in\mathcal{B}_i} w(S)U_i(S,\phi^S,e^S)$, we need to prove the following inequality for each $i\in N$:

$$f\left(\sum_{j\in N}(\lambda_j\bar{e}_j)\right) + \bar{\phi}_i - c(\bar{e}_i) \ge \sum_{S\in\mathcal{B}_i} w(S)\left[f\left(\sum_{j\in S}\lambda_j e_j^S\right) + \phi_i\left(f\left(\sum_{j\in S}\lambda_j e_j^S\right)\right) - c\left(e_i^S\right)\right]. \tag{4}$$

First, rearranging the terms, we obtain

$$f\left(\sum_{j\in N}(\lambda_j\bar{e}_j)\right) = f\left(\sum_{j\in N}\left(\lambda_j\sum_{S\in\mathcal{B}_j} w(S)e_j^S\right)\right) = f\left(\sum_{S\in\mathcal{B}} w(S)\left(\sum_{j\in S}\lambda_j e_j^S\right)\right).$$

The concavity of f yields:

$$f\left(\sum_{S\in\mathcal{B}_i} w(S)\left(\sum_{j\in S}\lambda_j e_j^S\right)\right) \ge \sum_{S\in\mathcal{B}_i} w(S)f\left(\sum_{j\in S}\lambda_j e_j^S\right)$$

for all $i\in N$.

By combining the last two expressions with f is an increasing function, one obtains that $f\left(\sum_{j\in N}(\lambda_j\bar{e}_j)\right) > \sum_{S\in\mathcal{B}_i} w(S)f\left(\sum_{j\in S}\lambda_j e_j^S\right)$ for all $i\in N$.

Secondly, $\bar{\phi}_i = \sum_{S\in\mathcal{B}_i} w(S)\phi_i^S\left(f\left(\sum_{j\in S}\lambda_j e_j^S\right)\right)$.

Lastly, because $c(\cdot)$ is convex, $c\left(\sum_{S\in\mathcal{B}_i} w(S)e_i^S\right) \le \sum_{S\in\mathcal{B}_i} w(S)c\left(e_i^S\right)$.

Therefore, we have proved inequality 4.

Part 2. In our team formation problem without transfers, we set $V^S \equiv \{u^S\in\mathbb{R}^{|S|} : \exists e^S \text{ s.t. } u_i^S \le U_i(S,e^S)\}$. It can be easily shown that V^S is closed and $v^S\in V^S$ if $v^S\le u^S$ for some $u^S\in V^S$, so the team formation problem without transfers is a cooperative game.

Claim. team formation problem without transfers is a balanced game.

Proof of the Claim. We need to prove that $u\in\mathbb{R}^N$ is V^N if, for any balanced set $\mathcal{B}$, $\pi^S(u)\in V^S$ for all $S\in\mathcal{B}$. Fix a balanced set $\mathcal{B}$. Let w be a corresponding balancing weight to $\mathcal{B}$. Consider any u such that $\pi^S(u)\in V^S$ for all $S\in\mathcal{B}$. This means that there exists e^S, such that $U_i(S,e^S) > u_i$

for all $S \in \mathcal{B}$ and for all $i \in S$. Since $\sum_{S \in \mathcal{B}_i} w(S) = 1$ for all $i \in N$, $\sum_{S \in \mathcal{B}_i} w(S) U_i(S, e^S) \geq u_i$.

Consider $\bar{e} \in \mathbb{R}^n$ where $\bar{e}_i = \sum_{S \in \mathcal{B}_i} w(S) e_i^S$. To prove that $u \in V^N$, we will now show that $U_i(N, \bar{e}) \geq \sum_{S \in \mathcal{B}_i} w(S) U_i(S, e^S)$. In other words, we need to show

$$f\left(\sum_{j \in N} (\lambda_j \bar{e}_j)\right) - c(\bar{e}_i) \geq \sum_{S \in \mathcal{B}_i} w(S) \left(f\left(\sum_{j \in S} \lambda_j e_j^S\right) - c(e_i^S)\right). \tag{5}$$

By rearranging the terms we get

$$f\left(\sum_{j \in N} (\lambda_j \bar{e}_j)\right) = f\left(\sum_{j \in N}\left(\lambda_j \sum_{S \in \mathcal{B}_j} w(S) e_j^S\right)\right) = f\left(\sum_{S \in \mathcal{B}} w(S)\left(\sum_{j \in S} \lambda_j e_j^S\right)\right).$$

On the other hand, the concavity of f yields

$$f\left(\sum_{S \in \mathcal{B}_i} w(S)\left(\sum_{j \in S} \lambda_j e_j^S\right)\right) \geq \sum_{S \in \mathcal{B}_i} w(S) f\left(\sum_{j \in S} \lambda_j e_j^S\right)$$

for all $i \in N$.

Combining the last two expressions with f being an increasing function, one obtains that $f\left(\sum_{j \in N} (\lambda_j \bar{e}_j)\right) > \sum_{S \in \mathcal{B}_i} w(S) f\left(\sum_{j \in S} \lambda_j e_j^S\right)$ for all $i \in N$.

Since $c(\cdot)$ is convex, $c\left(\sum_{S \in \mathcal{B}_i} w(S) e_i^S\right) < \sum_{S \in \mathcal{B}_i} w(S) c\left(e_i^S\right)$.

Therefore, we have proved the equation 5. $\square$

Proof of Lemma 5.

1. Let $e^S \in NE(S)$. Then $e_i^S = \arg\max_{e_i} U_i\left(S, (e_i, e_{-i}^S)\right)$ for all $i \in S$. The first order condition yields $\rho(|S|) f'\left(\sum_{j \in S} \lambda_j e_j^S\right) = \frac{a}{\lambda_i} e_i^S$ for all $i \in S$. Since the left hand side of the FOC is the same for all players at an equilibrium, $e_j^S = \frac{\lambda_j}{\lambda_i} e_i^S$ for all $j \in S$. By substituting e_j^S-s back into the FOC and rearranging the terms, we obtain $\rho(|S|) f'\left(\frac{a}{\lambda_i} t^S e_i^S\right) = \frac{a}{\lambda_i} e_i^S$. Since $e_\lambda(S) = \frac{a}{\lambda_i} t^S e_i^S$, we obtain equation 2. The uniqueness of Nash equilibria is proved by observing that equation 2 has a unique solution and $e_i^S = \frac{\lambda_i}{at(S)} e_\lambda(S)$.
   Equality 3 follows by substituting the expression $e_i^S = \frac{\lambda_i}{at(S)} e_\lambda(S)$ into utility function.

2. From part (a), the equilibrium weighted effort $e_\lambda(S)$ depends only on t^S and $\rho(|S|)$. This implies that the players' equilibrium utility from team S depends only on $\rho(|S|)$, t_i and t_{-i}. Therefore, we use the following notations: $e_\lambda(t, \rho)$ is the solution to the equation $e = \rho t f'(e)$ and $U_i(t_i, t_{-i}, \rho) = \rho f(e_\lambda(t_i + t_{-i}, \rho)) - \frac{t_i}{2}\left(\frac{e_\lambda(t_i + t_{-i}, \rho)}{t_i + t_{-i}}\right)$.

Because $|S| = |T|$, it suffices to show that $\frac{\partial U_i(t_i, t_{-i}, \rho)}{\partial t_{-i}} > 0$ for any i. Let $t = t_i + t_{-i}$. Then

$$\frac{\partial U_i(t_i, t_{-i}, \rho)}{\partial t_{-i}} = \rho f'(e_\lambda(t, \rho)) \frac{\partial e_\lambda(t, \rho)}{\partial t} - t_i \frac{e_\lambda(t, \rho)}{t} \frac{t \frac{\partial e_\lambda(t, \rho)}{\partial t} - e_\lambda(t, \rho)}{(t)^2}.$$

Since $e_\lambda(t, \rho) = \rho t f'(e_\lambda(t, \rho))$, using the implicit function theorem, we obtain

$$\frac{\partial e_\lambda(t, \rho)}{\partial t} = \frac{\rho f'\left(e_\lambda(t, \rho)\right)}{1 - \rho t f''\left(e_\lambda(t, \rho)\right)} = \frac{e_\lambda(t, \rho)}{\left(1 - \rho t f''(e_\lambda(t, \rho))\right) t} \tag{6}$$

where the second equality follows from the FOC. Now by substituting the expression for $\frac{\partial e_\lambda(t, \rho)}{\partial t}$ and rearranging the terms, we obtain

$$\frac{\partial U_i(t_i, t_{-i}, \rho)}{\partial t_{-i}} = \frac{1 - \rho t_i f''\left(e_\lambda(t, \rho)\right)}{1 - \rho t f''\left(e_\lambda(t, \rho)\right)} \left(\frac{e_\lambda(t, \rho)}{t}\right)^2 > 0 \tag{7}$$

as $f'' \leq 0$.

□

Lemma 6. *At equilibrium, each player's utility*

$$U_i(S) > \left(t(S) - \frac{t_i}{2}\right) \left(\frac{e_\lambda(S)}{t(S)}\right)^2$$

Proof. Lemma 5.a and the fundamental theorem of calculus yield,

$$\begin{aligned} U_i(S) &= \rho(|S|) f\left(e_\lambda(S)\right) - \frac{t_i}{2} \left(\frac{e_\lambda(S)}{t(S)}\right)^2 \\ &= \rho(|S|) \left(\int_0^{e_\lambda(S)} f'(e) de\right) - \frac{t_i}{2} \left(\frac{e_\lambda(S)}{t(S)}\right)^2. \end{aligned}$$

Since $f'' \leq 0$, for any $e \in [0, e_\lambda(S)]$, $f'(e) \geq f'(e_\lambda(S))$. Hence,

$$\begin{aligned} U_i(S) &> \rho(|S|)\left(\int_0^{e_\lambda(S)} f'(e_\lambda(S))de\right) - \frac{t_i}{2}\left(\frac{e_\lambda(S)}{t(S)}\right)^2 \\ &= e_\lambda(S)\rho(|S|)f'(e_\lambda(S)) - \frac{t_i}{2}\left(\frac{e_\lambda(S)}{t(S)}\right)^2. \end{aligned}$$

By substituting $\rho(|S|)f'(e_\lambda(S)) = \frac{e_\lambda(S)}{t(S)}$ (lemma 5.a), we obtain the desired result. $\square$

Lemma 7. *If $S \subseteq T$, then $\frac{e_\lambda(S)}{t(S)} \geq \frac{e_\lambda(T)}{t(T)}$*

Proof. Let us consider $\frac{\partial e_\lambda(t,\rho)/t}{\partial t} = \frac{\rho f''(e_\lambda(t,\rho))}{1-\rho t f''(e_\lambda(t,\rho))}\frac{e_\lambda(t,\rho)}{t} < 0$. This means $\frac{e_\lambda(t^S,\rho(|S|))}{t^S} \geq \frac{e_\lambda(t^T,\rho(|S|))}{t^T}$ since $t^T \geq t^S$. Now showing $e_\lambda(t^T,\rho(|T|)) \leq e_\lambda(t^T,\rho(|S|))$ completes the proof. Since $\rho(|T|) \leq \rho(|S|)$, it suffices to show $\frac{\partial e_\lambda(t,\rho)}{\partial \rho} > 0$. Because $e(t,\rho) = \rho t f'(e(t,\rho))$, by using the implicit function, we obtain

$$\frac{\partial e_\lambda(t,\rho)}{\partial \rho} = \frac{e_\lambda(t,\rho)}{\big(1-\rho t f''(e_\lambda(t,\rho))\big)\rho}. \tag{8}$$

This expression is positive as $f'' \leq 0$. $\square$

Proof of Proposition 5. Before we start the proof let us consider what happens to the equilibrium utility of player i if t_i changes.

$$\frac{\partial U_i(t_i,t_{-i},\rho)}{\partial t_i} = \rho f'\left(e_\lambda(t,\rho)\right)\frac{\partial e_\lambda(t,\rho)}{\partial t} - \frac{1}{2}\left(\frac{e_\lambda(t,\rho)}{t}\right)^2 - t_i\frac{e_\lambda(t,\rho)}{t}\frac{t\frac{\partial e_\lambda(t,\rho)}{\partial t} - e_\lambda(t,\rho)}{t^2}.$$

By using the expression for $\frac{\partial e_\lambda(t,\rho)}{\partial t}$ (equation 6) and rearranging the terms,

$$\frac{\partial U_i(t_i,t_{-i},\rho)}{\partial t_i} = \frac{1-(t_i-t_{-i})\rho f''(e_\lambda(t,\rho))}{2\left(1-\rho t f''(e_\lambda(t,\rho))\right)}\left(\frac{e_\lambda(t,\rho)}{t}\right)^2 \tag{9}$$

Now we are ready to start proving the first part of the proposition.

Part 1. There exists a stable partition if $r \leq 4$.

Consider partition Π_r which is the result of the algorithm used in the proof of proposition 4.

Claim. The above canonical partition is stable when $r \leq 4$.

Proof of the claim. To prove this claim, we first prove that no member of $\Pi_r(1)$ is a part of a blocking team. Thereafter, by iterating the argument to the other teams of Π_r in the order they were formed when the algorithm is applied, we complete the proof.

Observe that the algorithm and lemma 5.b imply that $\Pi_r(1) = \{1, \cdots, |\Pi_r(1)|\}$. To prove that no member of $\Pi_r(1)$ is a part of a blocking team, we proceed in several steps .

Step 1. No member of $\Pi_r(1)$ is a part of a blocking team with a size no less than $|\Pi_r(1)|$.

Proof of Step 1. Contrary to the claim in step 1, let S be a team that blocks Π_r and contains player $i \in \Pi_r(1)$ and $|S| > |\Pi_r(1)|$. If S contains 2 or more players of $\Pi_r(1)$, then let i be the lowest indexed one of these players. In addition, $i \neq 1$ because player 1 picked $\Pi_r(1)$ to match with. Therefore, step 1 is proved if $|\Pi_r(1)| = 1$, leaving the cases in which $|\Pi_r(1)| > 1$. Consider team $\{1, \cdots, |S|\}$, which must contain player i because $|S| > |\Pi_r(1)|$. By lemma 5.b, $U_i(\{1, \cdots, |S|\}) \geq U_i(S)$ because $t_1 + \cdots + t_{|S|} \geq t(S)$ as $1 \notin S$. In addition,

$$\begin{aligned} & U_j\left(\{1, \cdots, |S|\}\right) - U_j\left(\Pi_r(1)\right) = \\ & \underbrace{\rho(|S|) f\left(e_\lambda\left(\{1, \cdots, |S|\}\right)\right) - \rho(|\Pi_r(1)|) f\left(e_\lambda(\Pi_r(1))\right)}_{1} + \\ & \underbrace{\frac{t_j}{2}}_{2} \underbrace{\left(\left(\frac{e_\lambda(\Pi_r(1))}{t^{\Pi_r(1)}}\right)^2 - \left(\frac{e_\lambda\left(\{1, \cdots, |S|\}\right)}{\sum_{j=1}^{|S|} t_j}\right)^2\right)}_{3} \end{aligned} \tag{10}$$

where $j = 1$ or i. Observe here that expression 10 for player 1 and i differs in the second term. Term 3 is positive due to lemma 7. In addition, expression 10 is not positive for player 1 since she chose to match with $\Pi_r(1)$. Thus, the first term, common to both players, is not positive. Since $t_1 \geq t_i$, expression 10 is clearly negative for player i. This contradicts that i is a part of blocking team S, completing the proof of step 1.

Step 2. No member of $\Pi_r(1)$ can be a member of a blocking team with size 1.

Proof of Step 2. On the contrary, let $i \in \Pi_r(1)$ be a player who blocks Π_r by herself. Thanks to the previous step, $|\Pi_r(1)| \geq 2$. Since player 1 chose to match with $\Pi_r(1)$, $i \neq 1$ and $U_1(\Pi_r(1)) > U_1(\{1\})$. In addition, $U_i(\Pi_r(1)) \geq U_1(\Pi_r(1))$ from equation 3 since $t_1 \geq t_i$. Because player i blocks Π_r, $U_i(\Pi_r(1)) < U_i(\{i\})$. When combined these inequalities imply $U_i(\{i\}) > U_1(\{1\})$. However, from equation 9, $\frac{\partial U_i(t_i,0,\rho)}{\partial t_i} = \frac{1}{2}\left(\frac{e_\lambda(t_i,\rho)}{t_i}\right)^2 > 0$, implying that $U_1(\{1\}) \geq U_i(\{i\})$. This contradicts $U_i(\Pi_r(1)) < U_i(\{i\})$. This completes the proof of step 2.

Step 3. No member of $\Pi_r(1)$ can be a member of a blocking team with size 2.

Proof of Step 3. On the contrary, let S be a team with a size of 2 that contains a player in $\Pi_r(1)$. Thanks to step 1, $|\Pi_r(1)| \geq 3$. Let i be the lowest indexed player in $\Pi_r(1) \cap S$. Hence, i is the most productive player in S. Since player 1 chose to match with $\Pi_r(1)$, $i \neq 1$ and $U_1(\Pi_r(1)) > U_1(\{1, i\})$. In addition, $U_i(\Pi_r(1)) \geq U_1(\Pi_r(1))$ from equation 3 since $t_1 \geq t_i$. Because team

S blocks Π_r, $U_i(\Pi_r(1)) < U_i(S)$. When combined these inequalities imply $U_i(S) > U_1(\{1,i\})$. Recall equation 9 and observe that $\frac{\partial U_i(t_i,t_{-i},\rho)}{\partial t_i} \geq 0$ if $t_i \geq t_{-i}$. Therefore, $U_1(\{1,i\}) \geq U_i(\{i,i\})$. [6] From lemma 5.b, $U_i(\{i,i\}) \geq U_i(S)$ because i is the most productive player in S. The previous 2 inequalities imply $U_1(\{1,i\}) \geq U_i(S)$ which is a contradiction. This completes the proof of step 3.

Step 4. No member of $\Pi_r(1)$ can be a member of a blocking team with size 3.
Proof of Step 4. On the contrary, let S be a team with a size of 3 that contains a player in $\Pi_r(1)$. Thanks to step 1, $|\Pi_r(1)| = 4$, meaning $\Pi_r(1) = \{1,2,3,4\}$. As noted earlier $i \neq 1$. Now we will show that players 2 or 3 are not in S.
Suppose otherwise and let the lowest indexed player in S be player $i = 2$ or 3. Let $T = \{1,2,3\}$. Because player 1 chose to match with $\Pi_r(1)$, $U_1(\Pi_r(1)) > U_1(T)$. In addition, $U_i(\Pi_r(1)) \geq U_1(\Pi_r(1))$ from equation 3 since $t_1 \geq t_i$. Because team S blocks Π_r, $U_i(\Pi_r(1)) < U_i(S)$. When combined, these inequalities imply the following 2 inequalities:

$$U_1(T) < U_1(\Pi_r(1)) \tag{11}$$
$$U_i(S) - U_1(T) > U_i(\Pi_r(1)) - U_1(\Pi_r(1)) \tag{12}$$

From expression 3, $U_i(\Pi_r(1)) - U_1(\Pi_r(1)) = \frac{t_1 - t_i}{2}\left(\frac{e_\lambda\left(\rho(4), t^{\Pi_r(1)}\right)}{t^{\Pi_r(1)}}\right)^2$ which is decreasing in $\rho(4)$ thanks to the proof of lemma 7. As the left side of expression 12 does not depend on $\rho(4)$, the lower $\rho(4)$ is, the easier inequality 12 is to satisfy. However, $\rho(4)$ affects inequality 11 since player 1's utility $U_1(\Pi_r(1))$ moves in the same direction with $\rho(4)$ as

$$\begin{aligned}
\frac{\partial U_i(t_i, t_{-i}, \rho)}{\partial \rho} &= f(e_\lambda(t,\rho)) + \rho f'(e_\lambda(t,\rho))\frac{\partial e_\lambda(t,\rho)}{\partial \rho} - t_i \frac{e_\lambda(t,\rho))}{t^2}\frac{\partial e_\lambda(t,\rho)}{\partial \rho} \\
&= f(e_\lambda(t,\rho)) + \frac{t_{-i}}{t}\frac{e_\lambda(t,\rho))}{t}\frac{\partial e_\lambda(t,\rho)}{\partial \rho} > 0.
\end{aligned}$$

Therefore, if inequalities 11 and 12 are satisfied for $\rho(4)$, then inequality 12 must be satisfied

[6] The RHS expression can be thought of as player i being matched with a player whose productivity is the same as her own.

for ρ^* which equates inequality 11, i.e.,

$$U_1(T) = U_1(\Pi_r(1)) =$$
$$= \rho^* f\left(e_\lambda\left(t^{\Pi_r(1)}, \rho^*\right)\right) - \frac{t_1}{2}\left(\frac{e_\lambda\left(t^{\Pi_r(1)}, \rho^*\right)}{t^{\Pi_r(1)}}\right)^2.$$

To complete the proof of the current step, it suffices to show that inequality 12 is always violated when $\rho(4) = \rho^*$. The rest of the proof proceeds to show that the right hand side of inequality 12 is greater than the left hand side of inequality 12 for $\rho(4) = \rho^*$.

First, let us calculate the RHS of inequality 12. Observe that since $t(\Pi_r(1)) > t(T)$, lemma 5.b implies that $\rho^* < \rho(3)$ as $\frac{\partial U_i(t_i, t_{-i}, \rho)}{\partial \rho} > 0$. Consider function $\bar{\rho} : [t_{-1}(T), t_{-1}(\Pi_r(1))] \to [0, 1]$ which satisfies $U_1(t_1, t_{-1}, \bar{\rho}(t_{-1})) = U_1(T)$. Observe that $\bar{\rho}(t_{-1}(T)) = \rho(3)$ and $\bar{\rho}(t_{-1}(\Pi_r(1))) = \rho^*$. Hence, $U_1(t_1, t_{-1}(T), \bar{\rho}(t_{-1}(T))) = U_1(t_1, t_{-1}(\Pi_r(1)), \bar{\rho}(t_{-1}(\Pi_r(1))))$.

By the fundamental theorem of calculus, the right hand side of inequality 12 is

$$U_i(\Pi_r(1)) - U_1(\Pi_r(1)) = \frac{t_1 - t_i}{2}\left(\frac{e_\lambda\left(t(\Pi_r(1)), \bar{\rho}(t_{-1}(\Pi_r(1)))\right)}{t(\Pi_r(1))}\right)^2$$
$$= \frac{t_1 - t_i}{2}\left(\frac{e_\lambda\left(t(T), \bar{\rho}(t_{-1}(T))\right)}{t(T)} + \int_{t_{-1}(T)}^{t_{-1}(\Pi_r(1))} \frac{\partial\left(e_\lambda\left(t, \bar{\rho}(t_{-1})\right)/t\right)}{\partial t_{-1}} dt_{-1}\right)^2.$$

Let us find $\frac{\partial(e_\lambda(t, \bar{\rho}(t_{-1}))/t)}{\partial t_{-1}}$. Using the implicit function theorem and substituting the expressions for $\frac{\partial U_1(t_1, t_{-1}, \bar{\rho})}{\partial t_{-1}}$ and $\frac{\partial U_1(t_1, t_{-i}, \bar{\rho})}{\partial \bar{\rho}}$,

$$\frac{\partial \bar{\rho}(t_{-1})}{\partial t_{-1}} = -\frac{\frac{\partial U_1(t_1, t_{-1}, \bar{\rho})}{\partial t_{-1}}}{\frac{\partial U_1(t_1, t_{-i}, \bar{\rho})}{\partial \bar{\rho}}} = -\frac{\bar{\rho}\left(1 - t_1 \bar{\rho} f''\right)\left(\frac{e_\lambda(t, \bar{\rho})}{t}\right)^2}{\left(1 - t\bar{\rho} f''\right)\bar{\rho} f + (t - t_1)\left(\frac{e_\lambda(t, \bar{\rho})}{t}\right)^2}.$$

With the help of the above expression, we find that

$$\frac{\partial\left(e_{\lambda}\left(t,\bar{\rho}(t_{-1})\right)/t\right)}{\partial t_{-1}}=\frac{t\left(\frac{\partial e_{\lambda}(t,\bar{\rho})}{\partial\bar{\rho}}\frac{\partial\bar{\rho}}{\partial t_{-1}}+\frac{\partial e_{\lambda}(t,\bar{\rho})}{\partial t_{-1}}\right)-e_{\lambda}\left(t,\bar{\rho}\right)}{t^{2}}$$

$$=\frac{\bar{\rho}^{2}ff''-\left(\frac{e_{\lambda}(t,\bar{\rho})}{t}\right)^{2}}{\left(1-t\bar{\rho}f''\right)\bar{\rho}f+\left(t-t_{1}\right)\left(\frac{e_{\lambda}(t,\bar{\rho})}{t}\right)^{2}}\frac{e_{\lambda}\left(t,\bar{\rho}\right)}{t}.$$

The above inequality is negative as $f''\leq 0$. In addition,

$$\frac{\partial^{2}\left(e_{\lambda}\left(t,\bar{\rho}(t_{-1})\right)/t\right)}{\partial t_{-1}\partial f''}=\frac{\left(\bar{\rho}f-t_{1}\left(\frac{e_{\lambda}(t,\bar{\rho})}{t}\right)^{2}\right)\bar{\rho}^{2}f}{\left(\left(1-t\bar{\rho}f''\right)\bar{\rho}f+\left(t-t_{1}\right)\left(\frac{e_{\lambda}(t,\bar{\rho})}{t}\right)^{2}\right)^{2}}\frac{e_{\lambda}\left(t,\bar{\rho}\right)}{t}$$

$$>0.$$

The last inequality is due to lemma 6 and $f''\leq 0$. Consequently, $\frac{\partial(e_{\lambda}(t,\bar{\rho}(t_{-1}))/t)}{\partial t_{-1}}>-\frac{1}{t}\frac{e_{\lambda}(t,\bar{\rho}(t))}{t}$ where the RHS is the value of $\frac{\partial(e_{\lambda}(t,\bar{\rho}(t_{-1}))/t)}{\partial t_{-1}}$ when $f''\to-\infty$. Moreover, thanks to lemma 7, for any $t_{-1}\in[t_{-1}(T),t_{-1}(\Pi_{r}(1))]$, $\frac{e_{\lambda}(t(T),\bar{\rho}(t_{-1}(T)))}{t_{-1}(T)}\geq\frac{e_{\lambda}(t,\bar{\rho}(_{-1}))}{t}$. Therefore, for any $t_{-1}\in[t_{-1}(T),t_{-1}(\Pi_{r}(1))]$, $\frac{\partial(e_{\lambda}(t,\bar{\rho}(t_{-1}))/t)}{\partial t_{-1}}>-\frac{1}{t(T)}\frac{e_{\lambda}(t(T),\bar{\rho}(t_{-1}(T)))}{t(T)}$.

As a result, we find that

$$U_{i}(\Pi_{r}(1))-U_{1}(\Pi_{r}(1))>\frac{t_{1}-t_{i}}{2}\left(1-\frac{t(\Pi_{r}(1))-t(T)}{t(T)}\right)^{2}\left(\frac{e_{\lambda}\left(t(T),\rho(3)\right)}{t(T)}\right)^{2}$$

$$>\frac{t_{1}-t_{i}}{2}\left(\frac{t_{1}+t_{2}}{t_{1}+t_{2}+t_{3}}\right)^{2}\left(\frac{e_{\lambda}\left(t(T),\rho(3)\right)}{t(T)}\right)^{2}$$

The last inequality is due to $t_{3}\geq t_{4}$. Now we turn our attention to the left hand side inequality 12. Clearly, $U_{i}(S)-U_{1}(T)=U_{i}(T)-U_{1}(T)-(U_{i}(T)-U_{i}(S))$. The fundamental theorem of calculus along with expression 9 yields

$$U_{i}(T)-U_{i}(S)=\int_{t(S)}^{t(T)}\frac{1-\rho(3)t_{i}f''}{1-\rho(3)tf''}\left(\frac{e(t,\rho(3))}{t}\right)^{2}dt>0$$

It can easily be shown that $\frac{1-\rho(3)t_i f''}{1-\rho(3)tf''}$ increases with f''. Consequently,

$$U_i(T) - U_i(S) > \int_{t(S)}^{t(T)} \frac{t_i}{t} \left(\frac{e(t, \rho(3))}{t} \right)^2 dt$$

where the RHS is the value of $U_i(T) - U_i(S)$ corresponding to $f'' \to -\infty$. In addition, $\frac{e(t^T, \rho(3))}{t^T} < \frac{e(t,\rho(3))}{t}$ for any $t \in [t(S), t(T)]$ (lemma 4). Hence, $U_i(T) - U_i(S) > \frac{(t(T)-t(S))t_i}{t(T)} \left(\frac{e(t(T),\rho(3))}{t(T)} \right)^2$. Therefore,

$$\begin{aligned} U_i(S) - U_1(T) &< \frac{t_1 - t_i}{2} \left(\frac{e(t(T), \rho(3))}{t(T)} \right)^2 - \frac{(t(T) - t(S))t_i}{t(T)} \left(\frac{e(t(T), \rho(3))}{t(T)} \right)^2 \\ &< \frac{t_1 - t_i}{2} \frac{t_1 + t_2 + t_3 - 2t_i}{t_1 + t_2 + t_3} \left(\frac{e(t(T), \rho(3))}{t(T)} \right)^2. \end{aligned}$$

The last inequality is obtained as $t(T) - t(S) \geq t_1 - t_i$.

Now we are ready to compare the left and right hand sides of inequality 12. Clearly, $U_i(\Pi_r(1)) - U_1(\Pi_r(1)) > U_i(S) - U_1(T)$ as $\left(\frac{t_1+t_2}{t_1+t_2+t_3} \right)^2 > \frac{t_1+t_2+t_3-2t_i}{t_1+t_2+t_3}$. This means that players 2 and 3 cannot be part of a blocking team.

Now let us show that player 4 cannot be a member of a blocking team to Π_r. It suffices to show that $U_4(t_4, \sum_{j=1}^3 t_j, \rho(4)) \geq U_4(t_4, 2t_4, \rho(3))$. Let us perturb the productivity coefficients of players 4, 5, and 6 so that $t'_4 = t'_5 = t'_6 = t_3$. When the canonical algorithm is applied to the set of players with perturbed productivity, $\Pi_r(1) = \{1, 2, 3, 4\})$ because $U_1(\{1, 2, 3, 4\})$ does not decrease as $t'_4 \geq t^4$ and $U_1(\{1, 2, 3\})$, $U_1(\{1, 2\})$ and $U_1(\{1\})$ are unchanged. As we already know, player 3 cannot be a member of any blocking team, so player 4 also cannot a member of any blocking team. Hence, $U_3(t_3, \sum_{j=1}^3 t_j, \rho(4)) > U_3(t_3, 2t_3, \rho(3))$. Now we consider $U_4(t_4, \sum_{j=1}^3 t_j, \rho(4))$ and $U_4(t_4, 2t_4, \rho(3))$.

Observe that $U_4(t_4, \sum_{j=1}^3 t_j, \rho(4)) = U_3(t_3, \sum_{j=1}^3 t_j, \rho(4)) + \int_{t_3}^{t_4} \frac{\partial U_i(t_i, \sum_{j=1}^3 t_j, \rho(4))}{\partial t_i} dt_i$ thanks to the fundamental theorem of calculus. Let $t = t_1 + t_2 + t_3 + t_i$. Thanks to expression 9,

$$\begin{aligned} U_4(\sum_{j=1}^3 t_j, \rho(4)) &= U_3(t_3, \sum_{j=1}^3 t_j, \rho(4)) - \frac{1}{2} \int_{t_4}^{t_3} \frac{1 - \rho(4)(2t_i - t)f''}{1 - \rho(4)tf''} \left(\frac{e(t, \rho(4))}{t} \right)^2 dt_i \\ &> U_3(t_3, \sum_{j=1}^3 t_j, \rho(4)) - \frac{t_3 - t_4}{2} \left(\frac{e_\lambda(t(\{1, 2, 3, 4\}), \rho(4))}{t(\{1, 2, 3, 4\})} \right)^2 \end{aligned}$$

The last inequality is obtained by setting $f'' = 0$ and $\frac{e(t,\rho(4))}{t} = \frac{e_\lambda(t(\{1,2,3,4\}),\rho(4))}{t(\{1,2,3,4\})}$ as $\frac{1-\rho(4)(2t_i-t)f''}{1-\rho(4)tf''}$

is increasing in f'' and $\frac{e_\lambda(t(\{1,2,3,4\}),\rho(4))}{t(\{1,2,3,4\})} > \frac{e_\lambda(t,\rho(4))}{t}$ for any $t \in (t(\{1,2,3,4\}), t(\{1,2,3,3\})]$ due to lemma 7.

On the other hand, $U_4(t_4, 2t_4, \rho(3)) = U_3(t_3, 2t_3, \rho(3)) + \int_{t_3}^{t_4} \frac{\partial U_i(x,2x,\rho(3))}{\partial x} dx$ thanks to the fundamental theorem of calculus. It can be easily calculated that

$$\frac{\partial U_i(x, 2x, \rho(3))}{\partial x} = \left(\frac{1}{2} + \frac{2}{1 - 3x\rho(3)f''}\right)\left(\frac{e(3x, \rho(3))}{3x}\right)^2$$

Hence,

$$U_4(t_4, 2t_4, \rho(3)) < U_3(t_3, 2t_3, \rho(3)) - \frac{t_3 - t_4}{2}\left(\frac{e(3t_3, \rho(3))}{3t_3}\right)^2.$$

The last inequality obtained by setting $f'' = -\infty$ and $\frac{e(3x,\rho(3))}{3x} = \frac{e(3t_3,\rho(3))}{3t_3}$ as $\frac{2}{1-3t_i\rho(3)f''} > 0$ and $\frac{e(3t_3,\rho(3))}{3t_3} < \frac{e(3x,\rho(3))}{3x}$ for any $x \in [t_4, t_3)$ (lemma 7).

To show $U_4(t_4, \sum_{j=1}^3 t_j, \rho(4)) > U_4(t_4, 2t_4, \rho(3))$, we only need to show

$$U_3(t_3, \sum_{j=1}^{3} t_j, \rho(4)) - \frac{t_3 - t_4}{2}\left(\frac{e_\lambda(t(\{1,2,3,4\}), \rho(4))}{t(\{1,2,3,4\})}\right)^2 \geq$$
$$U_3(t_3, 2t_3, \rho(3)) - \frac{t_3 - t_4}{2}\left(\frac{e(3t_3, \rho(3))}{3t_3}\right)^2.$$

As we noted earlier, $U_3(t_3, \sum_{j=1}^3 t_j, \rho(4)) \geq U_3(t_3, 2t_3, \rho(3))$. In addition, $\frac{e(3t_3,\rho(3))}{3t_3} \geq \frac{e(t(\{1,2,3,4\}),\rho(4))}{t(\{1,2,3,4\})}$ thanks to lemma 7 as $t(\{1,2,3,4\}) > 3t_4$. This completes the proof of step 4 as well as the proof of part 1.

Part 2 Any stable partition is assortative when $r \leq 4$.

Proof of Part 2 Consider any non-assortative partition $\mathcal{B}_r$ where $r \leq 4$. Since $\mathcal{B}_r$ is non-assortative, there must exist $S, T \in \mathcal{B}_r$ such that $\max_{i \in S} t_i > \min_{i \in T} t_i$ and $\max_{i \in S} t_i > \min_{i \in T} t_i$. Let $\bar{i}/\underline{i}$ and $\bar{j}/\underline{j}$ be the players with the highest/lowest productivity from team S and T, respectively. Without loss of generality, we assume $t_{\bar{i}} \geq t_{\bar{j}}$. Now we show that $\mathcal{B}_r$ is blocked by some team. Then there are three cases to consider.

1. $|S| < |T|$
2. $|S| = |T|$
3. $|S| > |T|$

1. If $U_{\bar{j}}(S \backslash \{\underline{i}\} \cup \{\bar{j}\}) \geq U_{\bar{j}}(T \backslash \{\underline{j}\} \cup \{\bar{i}\})$, then $S \backslash \{\underline{i}\} \cup \{\bar{j}\}$ blocks $\mathcal{B}_r$. To see this, observe that, for any $i \in S \backslash \{\underline{i}\}$, $U_i(S \backslash \{\underline{i}\} \cup \{\bar{j}\}) > U_i(S)$ thanks to lemma 5.b as $t_{\bar{l}} > t_{\underline{i}}$. In

addition, $U_{\bar{j}}(S\backslash\{\underline{i}\} \cup \{\bar{j}\}) \geq U_{\bar{j}}(T\backslash\{\underline{j}\} \cup \{\bar{i}\}) > U_{\bar{j}}(T)$ where the last inequality is due to lemma 5.b as $t_{\bar{i}} > t_{\underline{j}}$. As a result, all the members of $S\backslash\{\underline{i}\} \cup \{\bar{j}\}$ strictly prefer $S\backslash\{\underline{i}\} \cup \{\bar{j}\}$ over $\mathcal{B}_r$, hence, block $\mathcal{B}_r$.

If $U_{\bar{j}}(S\backslash\{\underline{i}\}\cup\{\bar{j}\}) < U_{\bar{j}}(T\backslash\{\underline{j}\}\cup\{\bar{i}\})$, then $T\backslash\{\underline{j}\}\cup\{\bar{i}\}$ blocks $\mathcal{B}_r$. To see this, observe that, for any $i \in T\backslash\{\underline{l}\}$, $U_i(T\backslash\{\underline{j}\} \cup \{\bar{i}\}) > U_i(T)$ thanks to lemma 5.b as $t_{\bar{i}} > t_{\underline{j}}$. In addition, $U_{\bar{i}}(S\backslash\{\underline{i}\} \cup \{\bar{j}\}) > U_{\bar{i}}(S)$ thanks to lemma 5.b as $t_{\bar{j}} > t_{\underline{i}}$. Furthermore, equation 10 implies that $U_{\bar{i}}(T\backslash\{\underline{j}\} \cup \{\bar{i}\}) > U_{\bar{i}}(S\backslash\{\underline{i}\} \cup \{\bar{j}\})$ because $U_{\bar{j}}(T\backslash\{\underline{j}\} \cup \{\bar{i}\}) > U_{\bar{j}}(S\backslash\{\underline{i}\} \cup \{\bar{j}\})$ and $t_{\bar{i}} \geq t_{\bar{j}}$. As a result, all the members of $T\backslash\{\underline{j}\} \cup \{\bar{i}\}$ strictly prefer $T\backslash\{\underline{j}\} \cup \{\bar{i}\}$ over $\mathcal{B}_r$, hence, block $\mathcal{B}_r$.

2. Let $\bar{S}$ be the team of the lowest indexed $|S|$ players in $S \cup T$. Observe that $t^{\bar{S}} > t^S$ because $\bar{j} \in \bar{S}$ as $t_{\bar{j}} > t_{\underline{i}}$. Similarly, $t^{\bar{S}} > t^T$ because $\bar{i} \in \bar{S}$ as $t_{\bar{i}} > t_{\underline{j}}$. Hence, for all $i \in \bar{S} \cap S$, $U_i(\bar{S}) > U_i(S)$ and, for all $i \in \bar{S} \cap T$, $U_i(\bar{S}) > U_i(T)$ thanks to lemma 5.b. Hence, $\bar{S}$ blocks $\mathcal{B}_r$.

3. Let $\bar{S}$ be the team of the lowest indexed $|S|$ players in $S \cup T$. Similarly, let $\bar{T}$ be the team of the lowest indexed $|T|$ players in $S \cup T$. Observe that $\bar{i} \in \bar{S}$ and $\bar{i} \in \bar{T}$. In addition, by construction, $\bar{T} \subset \bar{S}$. As noted in case 2, $t^{\bar{S}} > t^S$. Furthermore, $t^{\bar{T}} > t^T$ because $\bar{i} \in \bar{T}$ as $t_{\bar{i}} > t_{\underline{j}}$. Consequently, by lemma 5.b, $U_i(\bar{S}) > U_i(S)$ for all $i \in \bar{S} \cap S$ and $U_i(\bar{T}) > U_i(T)$ for all $i \in \bar{T} \cap T$. Now we show that $\bar{T}$ blocks $\mathcal{B}_r$ if $U_{\bar{i}}(\bar{T}) \geq U_{\bar{i}}(\bar{S})$ or $\bar{S}$ blocks $\mathcal{B}_r$ if $U_{\bar{i}}(\bar{T}) < U_{\bar{i}}(\bar{S})$.

   Let $U_{\bar{i}}(\bar{T}) \geq U_{\bar{i}}(\bar{S})$. We already know that if $i \in \bar{T} \cap T$, $U_i(\bar{T}) > U_i(T)$. If $i \in \bar{T} \cap S$, then equation 10 implies that $U_i(\bar{T}) \geq U_i(\bar{S})$ as player i is the most productive player in $\bar{T}$ and $U_{\bar{i}}(\bar{T}) \geq U_{\bar{i}}(\bar{S})$. In addition, we already know that $U_i(\bar{S}) > U_i(S)$ for all $i \in \bar{T} \cap S$. Therefore, if $i \in S \cap \bar{T}$, then $U_i(\bar{T}) \geq U_i(S)$. As a result, $\bar{T}$ blocks $\mathcal{B}_r$.

   Let $U_{\bar{i}}(\bar{T}) < U_{\bar{i}}(\bar{S})$. We already know that if $i \in \bar{S} \cap S$, then $U_i(\bar{S}) > U_i(S)$. Consider $i \in \bar{S} \cap T$. From the proof of part 1, one obtains that for all $i \in \bar{S} \cap T$, $U_i(\bar{S}) \geq U_i(\bar{T})$ because $\bar{i}$ is the most productive player in $\bar{S}$, $U_{\bar{i}}(\bar{S}) > U_{\bar{i}}(\bar{T})$ and $\bar{i} \notin T$. In addition, we already know that, for all $i \in \bar{S} \cap T$, $U_i(\bar{T}) > U_i(T)$. Hence, for any $i \in \bar{S} \cap T$, $U_i(\bar{S}) > U_i(T)$. Consequently, $\bar{T}$ blocks $\mathcal{B}_r$.

$\square$

Economics Working Paper

2010-18: Martin Paldam: A check of Maddison's gdp data. Benford's Law with some range problems

2010-19: Torben M. Andersen and Marias H. Gestsson: Longevity, Growth and Intergenerational Equity - The Deterministic Case

2011-01: Torben M. Andersen: Welfare State - The Scandinavian Model

2011-02: Torben M. Andersen: Collective risk sharing: The social safety net and employment

2011-03: Rune Vejlin: Optimal Unemployment Insurance: How Important is the Demand Side?

2011-04: Mikkel Nørlem Hermansen: Non-existence of Steady State Equilibrium in the Neoclassical Growth Model with a Longevity Trend

2011-05: John Kennes, Daniel Monte and Norovsambuu Tumennasan: The Daycare Assignment Problem

2011-06: Rene Saran and Norovsambuu Tumennasan: Whose Opinion Counts?

2011-07: Mark Strøm Kristoffersen: Liquidity Constraints and Fiscal Stabilization Policy

2011-08: Martin Paldam: The cycle of development in Africa. A story about the power of economic ideas

2011-09: Thorvardur Tjörvi Ólafsson, Ásgerdur Pétursdóttir and Karen Á. Vignisdóttir: Price setting in turbulent times. Survey evidence from Icelandic firms

2011-10: Torben M. Andersen: Social policies and activation in the Scandinavian welfare model: the case of Denmark

2011-11: Norovsambuu Tumennasan: To Err is Human: Implementation in Quantal Response Equilibria

2011-12: Norovsambuu Tumennasan: Do Contracts Help? A Team Formation Perspective