

TI 2014-086/II  
Tinbergen Institute Discussion Paper



# Transfers and Exchange-Stability in Two-Sided Matching Problems

*Emiliya Lazarova*<sup>1</sup>

*Peter Borm*<sup>2</sup>

*Arantza Estévez-Fernández*<sup>3</sup>

<sup>1</sup> *University of East Anglia, United Kingdom;*

<sup>2</sup> *Tilburg University, the Netherlands;*

<sup>3</sup> *Faculty of Economics and Business Administration, VU University Amsterdam, and Tinbergen Institute, the Netherlands.*



Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam  
Burg. Oudlaan 50  
3062 PA Rotterdam  
The Netherlands  
Tel.: +31(0)10 408 8900  
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 525 8579

# Transfers and Exchange-Stability in Two-Sided Matching Problems

Emiliya Lazarova<sup>1</sup>      Peter Borm<sup>2</sup>      Arantza Estévez-Fernández<sup>3</sup>

June 10, 2014

## Abstract

In this paper we consider one-to-many matching problems where the preferences of the agents involved are represented by monetary reward functions. We characterize Pareto optimal matchings by means of contractually exchange stability and matchings of maximum total reward by means of compensation exchange stability. To conclude, we show that in going from an initial matching to a matching of maximum total reward, one can always provide a compensation schedule that will be ex-post stable in the sense that there will be no subset of agents who can all by deviation obtain a higher reward. The proof of this result uses the fact that the core of an associated compensation matching game with constraints is nonempty.

**Keywords:** matching, Pareto optimal matching, contractually exchange stability, compensation stability, compensation schedule.

**JEL classification:** C78, C71, D60.

## 1 Introduction

Restructuring is an ongoing process in today's dynamic business world where companies are constantly looking for ways to enhance efficiency, to gain competitive edge, and to keep up with industry trends. A tension exists, however, between a job allocation that is best from an individual worker's or department's point of view and that of the organization as a whole. Thus, centralized restructuring that involves moving staff across different departments in an institution, potentially, leaves some staff worse-off. When a restructuring schedule is

---

<sup>1</sup>School of Economics, University of East Anglia, United Kingdom. Email: [E.Lazarova@uea.ac.uk](mailto:E.Lazarova@uea.ac.uk)

<sup>2</sup>Department of Econometrics and Operations Research, Tilburg University, Warandelaan 2, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Email: [P.E.M.Borm@uvt.nl](mailto:P.E.M.Borm@uvt.nl)

<sup>3</sup>Tinbergen Institute and Department of Econometrics and Operations Research, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. [arantza.estevezfernandez@vu.nl](mailto:arantza.estevezfernandez@vu.nl)

devised one should consider, therefore, the opportunity for an ex-post reorganization by some departments or workers that may undermine the process as a whole.

The focus of study is the class of Pareto optimal matchings of workers to departments within an organization, and, within this set, we pay particular attention to those matchings that maximize an organization's overall reward. We characterize the set of Pareto optimal matchings and the set of matchings that maximize the institution's total reward by means of two stability notions: contractual exchange stability and compensation exchange stability, respectively.

Contractual exchange stability requires that any deviation, i.e. workers changing their assigned department or departments exchanging subsets of workers, is approved by all affected parties. Underlying this stability notion is the assumption that an agent grants approval to a deviation only if his reward does not decrease as a result of the change. Our notion of contractual exchange stability is in line with the notion of contractual individual stability introduced by Bogomolnaia and Jackson (2002) in the context of hedonic coalition formation games.

Compensation exchange stability allows the deviating parties to compensate the agents that are made worse-off by the change. Thus, a deviation is viable only if the net increase in rewards of the deviating agents exceeds the total losses of those agents whose rewards decrease due to the deviation. Clearly, some deviations which are not allowed under the notion of contractual exchange stability (because they leave some agents worse-off) are possible under the notion of compensation exchange stability (because those agents can be compensated).

Our notions of stability are related to the notion of exchange stability in the context of one-to-one matching problems developed by Alcalde (1995). More recently Chechlárová (2002) and Chechlárová and Manlove (2005) emphasize the importance of preference completeness (i.e. whether all agents are acceptable partners to all other agents) and consistency (i.e. when each agent who finds another agent acceptable, is also found acceptable by the former) for the existence of this type of exchange-stable allocations. Our framework for exchange stability is different in the sense that we consider many-to-one matching problems with peer effects (i.e. the reward of a worker in a matching depends on the group of coworkers assigned to the same department) and, moreover, allow for monetary rewards.

We also address the question if, given an initial matching of workers to departments, there exists a compensation schedule such that a centralized restructuring towards a matching of maximum total reward of the organization will be ex-post stable, in the sense that there will be no subset of workers or departments who can all by deviation obtain a higher reward. We analyze this question by introducing a transferable utility cooperative game, called a compensation matching game, and by showing that each compensation matching game has a

nonempty core, which ensures that a compensation schedule as described above exists.

The paper is structured as follows. Section 2 provides notions on matchings for matching problems with monetary rewards and peer effects used throughout the paper. In Section 3 we characterize Pareto optimality of matchings by means of contractually exchange stability on the class of strict matching problems. In Section 4 we characterize matchings with maximum total reward by means of compensation exchange stability. Section 5 analyzes the existence of stable compensation schedules when an organization restructures the initial situation towards a matching of maximum total reward.

## 2 Basic Notions

We consider two distinct types of economic agents: workers and departments. For ease of notation, we use lower case letters to denote a generic member  $i$  within the set of workers  $\mathcal{N}$ , and upper case letters to denote a generic department  $H$  within the set of departments  $\mathcal{D}$ . A vector of natural numbers  $q = (q_H)_{H \in \mathcal{D}}$  represents the capacity of each department with  $q_H$  being the capacity of department  $H$  with  $q_H \geq 1$ . We assume that  $|\mathcal{N}| = \sum_{H \in \mathcal{D}} q_H$ , i.e. the total number of workers equals the total capacity available at the departments. An allocation of workers to departments forms a partition of the set of workers in  $\mathcal{N}$ ,  $P = \{P_H\}_{H \in \mathcal{D}}$ , such that the size of the partition elements is consistent with the departmental capacity, i.e.  $|P_H| = q_H$  for all  $H \in \mathcal{D}$ . The set of all such partitions is denoted by  $\mathcal{P}(\mathcal{N}, \mathcal{D}, q)$ . If no confusion arises, we write  $\mathcal{P}$  instead of  $\mathcal{P}(\mathcal{N}, \mathcal{D}, q)$ .

An allocation of workers to departments is given by a *matching*  $\mu$  where  $\mu : \mathcal{N} \rightarrow \mathcal{D}$  such that<sup>4</sup>  $P(\mu) = \{\mu^{-1}(H)\}_{H \in \mathcal{D}} \in \mathcal{P}$ . The class of matchings is denoted by  $\mathcal{M}(\mathcal{N}, \mathcal{D}, q)$ . If no confusion arises, we write  $\mathcal{M}$  instead of  $\mathcal{M}(\mathcal{N}, \mathcal{D}, q)$ .

We assume that for each worker a matching leads to a monetary reward which depends on the department to which he is assigned and the identity of his peers in that department. Thus, for each  $i \in \mathcal{N}$  there is a monetary reward function  $r_i$  where for each  $(\mathcal{S}_i, H)$  with  $H \in \mathcal{D}$  and  $\mathcal{S}_i \in 2^{\mathcal{N} \setminus \{i\}}$  such that  $|\mathcal{S}_i| = q_H - 1$ ,  $r_i(\mathcal{S}_i, H)$  can be read as the reward that worker  $i$  receives when matched to department  $H$  with a set of coworkers  $\mathcal{S}_i \subseteq \mathcal{N} \setminus \{i\}$ .

Similarly, we assume that for each department  $H \in \mathcal{D}$  there is a monetary reward function  $\pi_H : \{\mathcal{S} \subset \mathcal{N} \mid |\mathcal{S}| = q_H\} \rightarrow \mathbb{R}$ .

---

<sup>4</sup>Notice that in our definition of a matching we do not allow for workers to remain unmatched and for departments to have unfilled capacity. The choice of our assumption is driven by the particular environment that we have in mind where existing workers are allocated to existing departments and firing or a closure of a department are not a possibility. Chechlárová and Manlove (2005) study a model with similar characteristics.

Given a matching  $\mu \in \mathcal{M}$ , we will use the shorthand notation  $r_i(\mu)$  to denote the monetary reward obtained by worker  $i \in \mathcal{N}$  when the department to which  $i$  belongs and the set of  $i$ 's coworkers are given by matching  $\mu$ , i.e.  $r_i(\mu) = r_i(\mu^{-1}(\mu(i)) \setminus \{i\}, \mu(i))$ . Similarly,  $\pi_H(\mu)$  denotes the reward obtained by department  $H \in \mathcal{D}$  from the set of workers assigned by  $\mu$  to  $H$ , i.e.  $\pi_H(\mu) = \pi_H(\mu^{-1}(H))$ .

The tuple  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  with  $r = (r_i)_{i \in \mathcal{N}}$  and  $\pi = (\pi_H)_{H \in \mathcal{D}}$  as above, defines a many-to-one matching problem with peer effects and monetary evaluations or simply *matching problem*.

Strict preference profiles are a common assumption in many matching models (cf. Gale and Shapley (1962)). Recently, some authors have investigated the link between indifference and inefficiencies, e.g. Erdil and Ergin (2008). In our setting, a *matching problem* is called *strict* if

- (i) for every  $i \in \mathcal{N}$ , all  $H_1, H_2 \in \mathcal{D}$ , and all  $S_1, S_2 \subseteq \mathcal{N}$  with  $|S_1| = q_{H_1}$ ,  $|S_2| = q_{H_2}$  and  $i \in S_1 \cap S_2$ ,  $r_i(S_1 \setminus \{i\}, H_1) \neq r_i(S_2 \setminus \{i\}, H_2)$  if  $S_1 \neq S_2$  or  $H_1 \neq H_2$ ; and
- (ii) for every  $H \in \mathcal{D}$ , all  $S_1, S_2 \subseteq \mathcal{N}$  with  $|S_1| = |S_2| = q_H$ ,  $\pi_H(S_1) \neq \pi_H(S_2)$  if  $S_1 \neq S_2$ .

### 3 Pareto Optimal Matchings

In a Pareto optimal matching it is not possible to make any agent better-off by changing their department, in the case of a worker, or their workers, in the case of a department, without making another agent worse-off. The formal definition of a Pareto optimal matching is given below.

**Definition 3.1** *Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem. A matching  $\mu \in \mathcal{M}$  is Pareto optimal if there is no other matching  $\mu'$  such that*

- $r_i(\mu') \geq r_i(\mu)$  for all  $i \in \mathcal{N}$ ,
- $\pi_H(\mu') \geq \pi_H(\mu)$  for all  $H \in \mathcal{D}$ , and
- *there exists a worker  $i \in \mathcal{N}$  such that  $r_i(\mu') > r_i(\mu)$  or a department  $H \in \mathcal{D}$  such that  $\pi_H(\mu') > \pi_H(\mu)$ .*

We will characterize Pareto optimality of matchings by means of a stability notion called *contractual exchange stability*. Recall that in our set up departments have fixed capacities and the number of available workers is consistent with the available capacities. Thus, it seems natural to assume that an agents' only possibility for deviation is to perform some switch with other agents on the same side (i.e. workers can exchange departments with other workers and

departments can exchange workers with other departments) provided that the exchange meets certain contractual criteria. In particular, if worker  $i$  prefers the department and coworkers of worker  $j$  in matching  $\mu$ , and vice versa, then  $i$  and  $j$  can exchange their places provided that their respective departments and peers are not made worse-off by the swap. Similarly, if two departments,  $G$  and  $H$ , can generate higher rewards by exchanging a (sub-)set of their workers under  $\mu$ , then, they can perform the swap provided that none of the involved workers is made worse-off. The description above is limited to two-way exchanges, but we assume that three-way or even more complicated exchanges can be performed as long as the affected parties (peers or departments) do not earn a lower reward as a result of the exchange.

**Definition 3.2** *Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem. A matching  $\mu \in \mathcal{M}$  is contractually exchange blocked via another matching  $\mu'$  if either there exists a blocking coalition  $\mathcal{S} \subseteq \mathcal{N}$  with  $|\mathcal{S}| \geq 2$  such that*

(i) *for each  $i \in \mathcal{S}$  there is a  $j \in \mathcal{S}$  with  $\mu(i) \neq \mu(j)$  such that  $\mu'(i) = \mu(j)$ ;*

(ii) *for each  $i \in \mathcal{S}$ ,  $r_i(\mu') > r_i(\mu)$ ;*

(iii) *for all  $j \in \mathcal{N} \setminus \mathcal{S}$ ,  $r_j(\mu') \geq r_j(\mu)$ ;*

*for all  $H \in \mathcal{D}$ ,  $\pi_H(\mu') \geq \pi_H(\mu)$ ;*

*or there there exists a blocking coalition  $\mathcal{H} \subseteq \mathcal{D}$  with  $|\mathcal{H}| \geq 2$  such that*

(i') *for each  $H \in \mathcal{H}$  it holds that  $(\mu')^{-1}(H) \neq \mu^{-1}(H)$  and  $(\mu')^{-1}(H) \subseteq \cup_{F \in \mathcal{H}} \mu^{-1}(F)$ ;*

(ii') *for each  $H \in \mathcal{H}$ ,  $\pi_H(\mu') > \pi_H(\mu)$ ;*

(iii') *for all  $i \in \mathcal{N}$ ,  $r_i(\mu') \geq r_i(\mu)$ ;*

*for all  $F \in \mathcal{D} \setminus \mathcal{H}$ ,  $\pi_F(\mu') \geq \pi_F(\mu)$ .*

In the above definition, it is required that the blocking coalition contains at least two agents of the same type who agree to a switch and these agents can earn a strictly higher reward as a result of the switch (requirements (ii) and (ii')). Requirements (iii) and (iii') of the definition, on the other hand, are important only for those agents who are affected by the switches proposed by the blocking coalition. For all other agents, this requirement is automatically satisfied.

Below we introduce the notion of a contractually exchange stable matching.

**Definition 3.3** *A matching  $\mu$  is contractually exchange stable if it cannot be contractually exchange blocked.*



**Theorem 3.4** *Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a strict matching problem. A matching  $\mu \in \mathcal{M}$  is Pareto optimal if and only if it is contractually exchange stable.*

**Proof.** First, let  $\mu$  be a contractually exchange stable matching. We show that  $\mu$  is Pareto optimal by contradiction. Suppose that  $\mu$  is not Pareto optimal. Therefore, there exists a matching  $\mu'$  such that

- (1)  $r_i(\mu') \geq r_i(\mu)$  for all  $i \in \mathcal{N}$ ,
- (2)  $\pi_H(\mu') \geq \pi_H(\mu)$  for all  $H \in \mathcal{D}$ , and
- (3) there exists a worker  $i \in \mathcal{N}$  such that  $r_i(\mu') > r_i(\mu)$  or a department  $H \in \mathcal{D}$  such that  $\pi_H(\mu') > \pi_H(\mu)$ .

Let  $\mathcal{H} = \{H \in \mathcal{D} | (\mu')^{-1}(H) \neq \mu^{-1}(H)\}$ . Note that  $|\mathcal{H}| \geq 2$  since  $\mu' \neq \mu$ . By definition of  $\mathcal{H}$ , we have that

- (i') for each  $H \in \mathcal{H}$ ,  $(\mu')^{-1}(H) \neq \mu^{-1}(H)$  and  $(\mu')^{-1}(H) \subseteq \cup_{F \in \mathcal{H}} \mu^{-1}(F)$ .

Since  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  is a strict matching problem and taking (2) above into account, we have that

- (ii') for each  $H \in \mathcal{H}$ ,  $\pi_H(\mu') > \pi_H(\mu)$ .

Moreover, by points (1) and (2) above, we know that

- (iii') for all  $i \in \mathcal{N}$ ,  $r_i(\mu') \geq r_i(\mu)$ ;  
for all  $F \in \mathcal{D} \setminus \mathcal{H}$ ,  $\pi_F(\mu') \geq \pi_F(\mu)$ .

Therefore,  $\mu$  can be blocked via  $\mu'$  by coalition  $\mathcal{H}$ . This establishes a contradiction to our premise that  $\mu$  is contractually exchange stable.

Next, let  $\mu$  be a Pareto optimal matching. We show that  $\mu$  is contractually exchange stable by contradiction. Suppose that  $\mu$  is not contractually exchange stable. Then, there exists a matching  $\mu'$  that contractually exchange blocks  $\mu$ . We consider two cases: (a)  $\mathcal{S}$  is a blocking coalition of workers and (b)  $\mathcal{H}$  is a blocking coalition of departments.

- (a) Let  $\mathcal{S} \subseteq \mathcal{N}$  with  $|\mathcal{S}| \geq 2$  be a blocking coalition of workers. From (ii) and (iii) in Definition 3.2, one readily derives that

- for all  $i \in \mathcal{N}$ ,  $r_i(\mu') \geq r_i(\mu)$  and
- for all  $H \in \mathcal{D}$ ,  $\pi_H(\mu') \geq \pi_H(\mu)$ .

Moreover, from (ii) in Definition 3.2 and  $\mathcal{S} \neq \emptyset$ , we find that

- $r_i(\mu') > r_i(\mu)$  for some  $i \in \mathcal{N}$ .

This establishes a contradiction to our premise that  $\mu$  is Pareto optimal.

(b) Let  $\mathcal{H} \subseteq \mathcal{D}$  with  $|\mathcal{H}| \geq 2$  be a blocking coalition of departments. From (ii') and (iii') in Definition 3.2, one readily derives that

- for all  $i \in \mathcal{N}$ ,  $r_i(\mu') \geq r_i(\mu)$  and
- for all  $H \in \mathcal{D}$ ,  $\pi_H(\mu') \geq \pi_H(\mu)$ .

Moreover, from (ii') in Definition 3.2 and  $\mathcal{H} \neq \emptyset$ , we find that

- $\pi_H(\mu') > \pi_H(\mu)$  for some  $H \in \mathcal{D}$ .

This establishes a contradiction to our premise that  $\mu$  is Pareto optimal. ■

The following example illustrates the need for restricting to the class of strict matching problems in Theorem 3.4.

**Example 3.5** Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem with  $\mathcal{N} = \{1, 2, 3\}$  and  $\mathcal{D} = \{A, B\}$ , capacities  $q_A = 1$  and  $q_B = 2$ ; and reward functions

$$\begin{aligned} r_1(\emptyset, A) &= 0, & r_1(\{2\}, B) &= 2, & r_1(\{3\}, B) &= 0, \\ r_2(\emptyset, A) &= 1, & r_2(\{1\}, B) &= 2, & r_2(\{3\}, B) &= 0, \\ r_3(\emptyset, A) &= 0, & r_3(\{1\}, B) &= 2, & r_3(\{2\}, B) &= 1; \end{aligned}$$

and

$$\begin{aligned} \pi_A(\{1\}) &= 1, & \pi_A(\{2\}) &= 1, & \pi_A(\{3\}) &= 3, \\ \pi_B(\{1, 2\}) &= 3, & \pi_B(\{1, 3\}) &= 2, & \pi_B(\{2, 3\}) &= 1. \end{aligned}$$

Clearly,  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  is not a strict matching problem since worker 1 has the same reward when he is assigned to department  $A$  and when he is assigned to department  $B$  with co-worker 3. Also, department  $A$  has the same reward when assigned to workers 1 and 2. We show that in this situation it is possible for a matching to be contractually exchange stable, but not Pareto optimal.

Consider the matching  $\mu_1$  such that  $\mu_1(1) = A$ ,  $\mu_1(2) = \mu_1(3) = B$ . This matching is contractually exchange stable. Any coalition of workers that attempts to block the matching must be either  $\{1, 2\}$ , or  $\{1, 3\}$ . First, consider workers 1 and 2: 1 does not want to exchange assignments with 2 because his rewards are the same under  $\mu_1$  and under a new matching in which he is matched to  $B$  together with 3. Furthermore, matching  $\mu_1$  cannot be blocked by workers 1 and 3 because for worker 3 we have that  $r_3(\emptyset, A) = 0 < 1 = r_3(\mu_1)$ .

Similarly, we can show that  $\mu_1$  cannot be contractually exchange blocked by  $\mathcal{H} = \{A, B\}$ . Departments  $A$  and  $B$  cannot block  $\mu_1$  by exchanging workers 1 and 3 because worker 3 earns lower reward when matched to department  $A$  than when matched with 2 to department  $B$ . Departments  $A$  and  $B$  cannot contractually exchange block  $\mu_1$  by exchanging workers 1 and 2 because department  $A$  earns the same rewards when assigned to workers 1 or 2.

However, the matching  $\mu_1$  is not Pareto optimal. To see that, consider matching  $\mu_2$  such that  $\mu_2(1) = \mu_2(3) = B$  and  $\mu_2(2) = A$  with  $r_1(\mu_2) = 0 = r_1(\mu_1)$ ,  $r_2(\mu_2) = 1 > 0 = r_2(\mu_1)$ ,  $r_3(\mu_2) = 2 > 1 = r_3(\mu_1)$ ,  $\pi_A(\mu_2) = 1 = \pi_A(\mu_1)$  and  $\pi_B(\mu_2) = 2 > 1 = \pi_B(\mu_1)$ .  $\blacklozenge$

Note that the requirement of a strict matching problem in Theorem 3.4 is only used to show that every contractually exchange stable matching is Pareto optimal. Therefore, we have the following result.

**Proposition 3.6** *Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem. If a matching  $\mu \in \mathcal{M}$  is Pareto optimal, then it is contractually exchange stable.*

## 4 Total Reward Maximizing Matchings

Pareto optimal matchings can be ranked on the basis of corresponding total joint rewards of all workers and departments. A central manager may thus be interested not only in achieving a Pareto optimal matching, but in achieving a Pareto optimal matching that corresponds to maximum total reward. Note that we will refer to such matchings as matchings of maximum total reward. We give a formal definition of this notion below.

**Definition 4.1** *Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem. A matching  $\mu \in \mathcal{M}$  is of maximum total reward if*

$$\sum_{i \in \mathcal{N}} r_i(\mu) + \sum_{H \in \mathcal{D}} \pi_H(\mu) \geq \sum_{i \in \mathcal{N}} r_i(\mu') + \sum_{H \in \mathcal{D}} \pi_H(\mu') \text{ for all } \mu' \in \mathcal{M}.$$

Note that every matching  $\mu$  of maximum total reward is automatically Pareto optimal.

**Example 4.2** Reconsider the matching problem of Example 3.5. In this problem, there are two Pareto optimal matchings:  $\mu_1$  such that  $\mu_1(2) = A$  and  $\mu_1(1) = \mu_1(3) = B$ ; and  $\mu_2$  such that  $\mu_2(3) = A$  and  $\mu_2(1) = \mu_2(2) = B$ . There is, however, only one matching with a maximum total reward and that is  $\mu_2$  as the total reward of  $\mu_1$  is 6 and that of  $\mu_2$  is 10.  $\blacklozenge$

Next, we present a new stability notion, compensation stability, to (later) characterize the set of matchings of maximum total reward.

**Definition 4.3** Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem and let  $\mu \in \mathcal{M}$  be a matching. The matching  $\mu$  is compensation stable if there is no other matching  $\mu'$  such that

$$\sum_{i \in \mathcal{S} \cup \mathcal{T}} r_i(\mu') + \sum_{H \in \mathcal{F}} \pi_H(\mu') > \sum_{i \in \mathcal{S} \cup \mathcal{T}} r_i(\mu) + \sum_{H \in \mathcal{F}} \pi_H(\mu),$$

where  $\mathcal{S} = \{i \in \mathcal{N} \mid \mu'(i) \neq \mu(i)\}$ ,  $\mathcal{T} = \{i \in \mathcal{N} \setminus \mathcal{S} \mid \mu^{-1}(\mu(i)) \cap \mathcal{S} \neq \emptyset\}$ , and  $\mathcal{F} = \{H \in \mathcal{D} \mid (\mu')^{-1}(H) \neq \mu^{-1}(H)\}$ , i.e.  $\mathcal{S}$  is the set of workers that are assigned to a different department,  $\mathcal{T}$  is the set of workers that do not change department but change peers, and  $\mathcal{F}$  is the group of departments that get a different set of workers.

**Theorem 4.4** Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem and let  $\mu \in \mathcal{M}$  be a matching. Then,  $\mu$  is of maximum total reward if and only if it is compensation stable.

**Proof.** First, let  $\mu$  be a compensation stable matching. We show that  $\mu$  is of maximum total reward by contradiction. Suppose that  $\mu$  is not of maximum total reward. Then, there is a matching  $\mu'$  such that

$$\sum_{i \in \mathcal{N}} r_i(\mu') + \sum_{H \in \mathcal{D}} \pi_H(\mu') > \sum_{i \in \mathcal{N}} r_i(\mu) + \sum_{H \in \mathcal{D}} \pi_H(\mu). \quad (1)$$

With  $\mathcal{S} = \{i \in \mathcal{N} \mid \mu'(i) \neq \mu(i)\}$ ,  $\mathcal{T} = \{i \in \mathcal{N} \setminus \mathcal{S} \mid \mu^{-1}(\mu(i)) \cap \mathcal{S} \neq \emptyset\}$ , and  $\mathcal{F} = \{H \in \mathcal{D} \mid (\mu')^{-1}(H) \neq \mu^{-1}(H)\}$ , it follows that

$$r_i(\mu') = r_i(\mu) \text{ for all } i \in \mathcal{N} \setminus (\mathcal{S} \cup \mathcal{T}) \text{ and} \quad (2)$$

$$\pi_H(\mu') = \pi_H(\mu) \text{ for all } H \in \mathcal{D} \setminus \mathcal{F}. \quad (3)$$

Then,

$$\begin{aligned} \sum_{i \in \mathcal{S} \cup \mathcal{T}} r_i(\mu') + \sum_{H \in \mathcal{F}} \pi_H(\mu') &= \sum_{i \in \mathcal{N}} r_i(\mu') + \sum_{H \in \mathcal{D}} \pi_H(\mu') - \sum_{i \in \mathcal{N} \setminus (\mathcal{S} \cup \mathcal{T})} r_i(\mu') - \sum_{H \in \mathcal{D} \setminus \mathcal{F}} \pi_H(\mu') \\ &= \sum_{i \in \mathcal{N}} r_i(\mu') + \sum_{H \in \mathcal{D}} \pi_H(\mu') - \sum_{i \in \mathcal{N} \setminus (\mathcal{S} \cup \mathcal{T})} r_i(\mu) - \sum_{H \in \mathcal{D} \setminus \mathcal{F}} \pi_H(\mu) \\ &> \sum_{i \in \mathcal{N}} r_i(\mu) + \sum_{H \in \mathcal{D}} \pi_H(\mu) - \sum_{i \in \mathcal{N} \setminus (\mathcal{S} \cup \mathcal{T})} r_i(\mu) - \sum_{H \in \mathcal{D} \setminus \mathcal{F}} \pi_H(\mu) \\ &= \sum_{i \in \mathcal{S} \cup \mathcal{T}} r_i(\mu) + \sum_{H \in \mathcal{F}} \pi_H(\mu) \end{aligned}$$

where the second equality follows by equations (2) and (3), and the inequality follows by equation (1). This establishes a contradiction to our premise that  $\mu$  is compensation stable.

Next, let  $\mu$  be a matching of maximum total reward. We show that  $\mu$  is compensation stable by contradiction. Suppose that  $\mu$  is not compensation stable. Then, there is a matching

$\mu'$  such that

$$\sum_{i \in \mathcal{T} \cup \mathcal{S}} r_i(\mu') + \sum_{H \in \mathcal{F}} \pi_H(\mu') > \sum_{i \in \mathcal{T} \cup \mathcal{S}} r_i(\mu) + \sum_{H \in \mathcal{F}} \pi_H(\mu), \quad (4)$$

where  $\mathcal{S} = \{i \in \mathcal{N} \mid \mu'(i) \neq \mu(i)\}$ ,  $\mathcal{T} = \{i \in \mathcal{N} \setminus \mathcal{S} \mid \mu^{-1}(\mu(i)) \cap \mathcal{S} \neq \emptyset\}$ , and  $\mathcal{F} = \{H \in \mathcal{D} \mid (\mu')^{-1}(H) \neq \mu^{-1}(H)\}$ . It follows that

$$\pi_H(\mu') = \pi_H(\mu) \text{ for all } H \in \mathcal{D} \setminus \mathcal{F} \text{ and} \quad (5)$$

$$r_i(\mu') = r_i(\mu) \text{ for all } i \in \mathcal{N} \setminus (\mathcal{S} \cup \mathcal{T}). \quad (6)$$

Then,

$$\begin{aligned} \sum_{i \in \mathcal{N}} r_i(\mu') + \sum_{H \in \mathcal{D}} \pi_H(\mu') &= \sum_{i \in \mathcal{T} \cup \mathcal{S}} r_i(\mu') + \sum_{H \in \mathcal{F}} \pi_H(\mu') + \sum_{i \in \mathcal{N} \setminus (\mathcal{S} \cup \mathcal{T})} r_i(\mu') + \sum_{H \in \mathcal{D} \setminus \mathcal{F}} \pi_H(\mu') \\ &= \sum_{i \in \mathcal{T} \cup \mathcal{S}} r_i(\mu') + \sum_{H \in \mathcal{F}} \pi_H(\mu') + \sum_{i \in \mathcal{N} \setminus (\mathcal{S} \cup \mathcal{T})} r_i(\mu) + \sum_{H \in \mathcal{D} \setminus \mathcal{F}} \pi_H(\mu) \\ &> \sum_{i \in \mathcal{T} \cup \mathcal{S}} r_i(\mu) + \sum_{H \in \mathcal{F}} \pi_H(\mu) + \sum_{i \in \mathcal{N} \setminus (\mathcal{S} \cup \mathcal{T})} r_i(\mu) + \sum_{H \in \mathcal{D} \setminus \mathcal{F}} \pi_H(\mu) \\ &= \sum_{i \in \mathcal{N}} r_i(\mu) + \sum_{H \in \mathcal{D}} \pi_H(\mu) \end{aligned}$$

where the second equality follows by equations (5) and (6), and the inequality follows by equation (4). This establishes a contradiction to our premise that  $\mu$  is of maximum total reward.  $\blacksquare$

## 5 From an Initial Matching to a Maximum Total Reward Matching

In this section, we analyze situations in which there is an initial matching of workers to departments which does not generate maximum total reward. Thus, there exists a possible restructuring where agents are reassigned by means of a matching of maximum total reward. The question that arises is how to compensate those workers or departments that are worse off in the new situation. This question will be analyzed using a cooperative matching game with transferable utility.

A *cooperative (transferable utility) game* in characteristic function form is a pair  $(N, v)$  where  $N$  is a finite set of players and  $v : 2^N \rightarrow \mathbb{R}$  satisfying  $v(\emptyset) = 0$ . In general,  $v(\mathcal{C})$  represents the maximal joint reward that coalition  $\mathcal{C} \in 2^N$  can obtain when its members cooperate in an optimal way, without help of the players in  $N \setminus \mathcal{C}$ .

The *core* of a game  $(N, v)$  is defined by

$$\text{Core}(v) = \left\{ x \in \mathbb{R}^N \mid \sum_{i \in N} x_i = v(N), \sum_{i \in \mathcal{C}} x_i \geq v(\mathcal{C}) \text{ for all } \mathcal{C} \in 2^N \right\},$$

i.e. the core is the set of efficient allocations of  $v(N)$  to which no coalition can reasonably object. An important subclass of games with nonempty core is the class of convex games (see Shapley (1971)). A game  $(N, v)$  is said to be *convex* if  $v(\mathcal{C}^1 \cup \{i\}) - v(\mathcal{C}^1) \leq v(\mathcal{C}^2 \cup \{i\}) - v(\mathcal{C}^2)$  for every  $i \in N$  and every  $\mathcal{C}^1 \subseteq \mathcal{C}^2 \subseteq N \setminus \{i\}$ .

Before starting with the formal analysis of compensation in matching problems with an initial matching, we provide an illustrative example.

**Example 5.1** Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem with  $\mathcal{N} = \{1, 2, 3, 4\}$  and  $\mathcal{D} = \{A, B\}$ , capacities  $q_A = 2$  and  $q_B = 2$ , and reward functions

$$\begin{aligned} r_1(\{2\}, A) &= r_1(\{2\}, B) = 20, & r_1(\{3\}, A) &= r_1(\{3\}, B) = 7, & r_1(\{4\}, A) &= r_1(\{4\}, B) = 14, \\ r_2(\{1\}, A) &= r_2(\{1\}, B) = 10, & r_2(\{3\}, A) &= r_2(\{3\}, B) = 3, & r_2(\{4\}, A) &= r_2(\{4\}, B) = 9, \\ r_3(\{1\}, A) &= r_3(\{1\}, B) = 15, & r_3(\{2\}, A) &= r_3(\{2\}, B) = 7, & r_3(\{4\}, A) &= r_3(\{4\}, B) = 10, \\ r_4(\{1\}, A) &= r_4(\{1\}, B) = 14, & r_4(\{2\}, A) &= r_4(\{2\}, B) = 19, & r_4(\{3\}, A) &= r_4(\{3\}, B) = 10 \end{aligned}$$

and

$$\pi_A(\mathcal{S}) = \pi_B(\mathcal{S}) = 10$$

for every  $\mathcal{S} \in 2^{\mathcal{N}}$  with  $|\mathcal{S}| = 2$ .

Assume that the initial matching  $\mu_0$  is given by  $\mu_0(1) = \mu_0(4) = A$  and  $\mu_0(2) = \mu_0(3) = B$ . In this case, the initial situation has a total reward of 58 with  $r_1(\mu_0) = 14$ ,  $r_2(\mu_0) = 3$ ,  $r_3(\mu_0) = 7$ ,  $r_4(\mu_0) = 14$ ,  $\pi_A(\mu_0) = 10$  and  $\pi_B(\mu_0) = 10$ , while the maximum social reward is 70 which can be achieved, for example, by  $\hat{\mu}$  with  $\hat{\mu}(1) = \hat{\mu}(2) = A$  and  $\hat{\mu}(3) = \hat{\mu}(4) = B$  with  $r_1(\hat{\mu}) = 20$ ,  $r_2(\hat{\mu}) = 10$ ,  $r_3(\hat{\mu}) = 10$ ,  $r_4(\hat{\mu}) = 10$ ,  $\pi_A(\hat{\mu}) = 10$  and  $\pi_B(\hat{\mu}) = 10$ . Note that  $r_4(\hat{\mu}) = 10 < r_4(\mu_0) = 14$ .  $\blacklozenge$

Given a matching problem  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  and an arbitrary initial matching  $\mu_0$ , we are interested in studying adequate compensations among the workers and departments when they come together to improve their joint situation by adopting a matching of maximum total reward. In order to do this, we define a corresponding cooperative game whose core elements provide stable allocations of the extra rewards obtained by means of cooperation.

Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem with constraints and let  $\mu_0$  be an initial matching. We denote by  $\mathcal{M}^{\max}$  the set of matchings of maximal total reward, i.e.

$$\mathcal{M}^{\max} = \{\hat{\mu} \in \mathcal{M} \mid \hat{\mu} \text{ is of maximal total reward}\}.$$

For  $i \in \mathcal{N}$  and  $\hat{\mu} \in \mathcal{M}^{\max}$ , we denote by  $d_i(\hat{\mu})$  the difference of reward for agent  $i \in \mathcal{N}$  when going from  $\mu_0$  to  $\hat{\mu}$ , i.e.  $d_i(\hat{\mu}) = r_i(\hat{\mu}) - r_i(\mu_0)$ ; we set  $d_i^+(\hat{\mu}) = \max\{0, r_i(\hat{\mu}) - r_i(\mu_0)\}$  and  $d_i^-(\hat{\mu}) = -\min\{0, r_i(\hat{\mu}) - r_i(\mu_0)\}$ . Note that for every  $\hat{\mu} \in \mathcal{M}^{\max}$  it follows  $d_i(\hat{\mu}) = d_i^+(\hat{\mu}) - d_i^-(\hat{\mu})$ . We analogously define  $d_H(\hat{\mu}) = \pi_H(\hat{\mu}) - \pi_H(\mu_0)$ , and set  $d_H^+(\hat{\mu}) = \max\{0, \pi_H(\hat{\mu}) - \pi_H(\mu_0)\}$ , and  $d_H^-(\hat{\mu}) = -\min\{0, \pi_H(\hat{\mu}) - \pi_H(\mu_0)\}$  for every  $H \in \mathcal{D}$  and every every  $\hat{\mu} \in \mathcal{M}^{\max}$ .

**Definition 5.2** Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem and let  $\mu_0$  be an initial matching. The corresponding compensation matching game,  $(\mathcal{N} \cup \mathcal{D}, v)$ , is defined by

$$v(\mathcal{C}) = \min_{\hat{\mu} \in \mathcal{M}^{\max}} \left\{ v^{\hat{\mu}}(\mathcal{C}) \right\} \quad (7)$$

for every  $\mathcal{C} \subset \mathcal{N} \cup \mathcal{D}$ , where

$$v^{\hat{\mu}}(\mathcal{C}) = \max \left\{ 0, \sum_{i \in \mathcal{C} \cap \mathcal{N}} d_i^+(\hat{\mu}) + \sum_{H \in \mathcal{C} \cap \mathcal{D}} d_H^+(\hat{\mu}) - \sum_{i \in \mathcal{N}} d_i^-(\hat{\mu}) - \sum_{H \in \mathcal{D}} d_H^-(\hat{\mu}) \right\}. \quad (8)$$

Note that given a matching with maximum total reward  $\hat{\mu}$ , we conservatively assume that a coalition  $\mathcal{C}$  is required to compensate *all* agents that suffer from changing the initial situation. Being conservative, it is also assumed that a most disadvantageous (from the perspective of  $\mathcal{C}$ ) matching of maximum total reward will be adopted. Note that  $v(\mathcal{N} \cup \mathcal{D})$  equals the increase in total rewards achievable by adopting any matching with maximum total reward.

The following example illustrates the computation of matching games with constraints.

**Example 5.3** Reconsider the matching problem  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  of Example 5.1 and let  $\mu_0$  be given by  $\mu_0(1) = \mu_0(4) = A$  and  $\mu_0(2) = \mu_0(3) = B$ , with a social reward of 58. There are four matchings with maximum total reward, namely

$$\begin{aligned} \hat{\mu}_1: \quad & \hat{\mu}_1(1) = \hat{\mu}_1(2) = A, \quad \hat{\mu}_1(3) = \hat{\mu}_1(4) = B, \\ \hat{\mu}_2: \quad & \hat{\mu}_2(3) = \hat{\mu}_2(4) = A, \quad \hat{\mu}_2(1) = \hat{\mu}_2(2) = B, \\ \hat{\mu}_3: \quad & \hat{\mu}_3(1) = \hat{\mu}_3(3) = A, \quad \hat{\mu}_3(2) = \hat{\mu}_3(4) = B, \\ \hat{\mu}_4: \quad & \hat{\mu}_4(2) = \hat{\mu}_4(4) = A, \quad \hat{\mu}_4(1) = \hat{\mu}_4(3) = B. \end{aligned}$$

The maximum total reward for this problem is 70, therefore there is an extra reward of 12 when going from the initial matching to one of maximum total reward.

The corresponding extra rewards for the institutions when going from the initial matching to a matching with maximum total reward is always 0 since they have constant reward functions. The individual extra rewards and losses for the agents when going from the initial matching to a matching of maximum total reward are provided in Table 1.

Since the departments are indifferent between all matching functions, we can omit them from the study of compensations and the player set  $\mathcal{N} \cup \mathcal{D}$  can be restricted to  $\mathcal{N}$  only.

$\hat{\mu}$	$d_1(\hat{\mu})$	$d_2(\hat{\mu})$	$d_3(\hat{\mu})$	$d_4(\hat{\mu})$
$\hat{\mu}_1$	6	7	3	-4
$\hat{\mu}_2$	6	7	3	-4
$\hat{\mu}_3$	-7	6	8	5
$\hat{\mu}_4$	-7	6	8	5

  

$\hat{\mu}$	$d_1^+(\hat{\mu})$	$d_2^+(\hat{\mu})$	$d_3^+(\hat{\mu})$	$d_4^+(\hat{\mu})$	$\hat{\mu}$	$d_1^-(\hat{\mu})$	$d_2^-(\hat{\mu})$	$d_3^-(\hat{\mu})$	$d_4^-(\hat{\mu})$
$\hat{\mu}_1$	6	7	3	0	$\hat{\mu}_1$	0	0	0	4
$\hat{\mu}_2$	6	7	3	0	$\hat{\mu}_2$	0	0	0	4
$\hat{\mu}_3$	0	6	8	5	$\hat{\mu}_3$	7	0	0	0
$\hat{\mu}_4$	0	6	8	5	$\hat{\mu}_4$	7	0	0	0

Table 1: Individual extra rewards and losses for the agents in Example 5.3.

The value of coalition  $\{1, 2, 3\}$  in the compensation matching game is explained below. All coalitional values are given in Table 2 .

$$\begin{aligned}
v(\{1, 2, 3\}) &= \min_{l \in \{1, 2, 3, 4\}} \left\{ v^{\hat{\mu}_l}(\{1, 2, 3\}) \right\} \\
&= \min_{l \in \{1, 2, 3, 4\}} \left\{ \max \left\{ 0, \sum_{i=1}^3 d_i^+(\hat{\mu}_l) - \sum_{i=1}^4 d_i^-(\hat{\mu}_l) \right\} \right\} \\
&= \min \{ \max \{0, 16 - 4\}, \max \{0, 16 - 4\}, \max \{0, 14 - 7\}, \max \{0, 14 - 7\} \} \\
&= \min \{12, 12, 7, 7\} \\
&= 7.
\end{aligned}$$

$\mathcal{C}$	$\{1\}$	$\{2\}$	$\{3\}$	$\{4\}$	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{2,3\}$	$\{2,4\}$	$\{3,4\}$	$\{1,2,3\}$	$\{1,2,4\}$	$\{1,3,4\}$	$\{2,3,4\}$	$\mathcal{N}$
$v(\mathcal{C})$	0	0	0	0	0	1	0	6	3	0	7	4	5	6	12

Table 2: Coalitional values in Example 5.3.

Note that the core of this game is not empty since, for instance,  $(0, 0, 7, 5) \in \text{Core}(v)$ . Here, the core element  $(0, 0, 7, 5)$  can be interpreted as follows: when choosing matching  $\hat{\mu}_3$ , worker 2 is helped by worker 3 to fully compensate worker 1 with 6 and 1, respectively.

Note that  $(\mathcal{N}, v)$  is not convex since for  $i = 4$ ,  $\mathcal{C}^1 = \{2\}$ , and  $\mathcal{C}^2 = \{2, 3\}$  we have

$$v(\mathcal{C}^1 \cup \{i\}) - v(\mathcal{C}^1) = 3 - 0 > 6 - 6 = v(\mathcal{C}^2 \cup \{i\}) - v(\mathcal{C}^2). \quad \blacklozenge$$

It turns out that each compensation matching game has a nonempty core. This can be shown using a relation between compensation matching games and bankruptcy games.



A *bankruptcy problem* is defined by a tuple  $(N, E, c)$  where  $N = \{1, \dots, n\}$  is the set of agents,  $E$  is the estate that must be shared among the agents, and  $c \in \mathbb{R}^N$  is the vector of claims of the agents satisfying  $\sum_{i \in N} c_i \geq E$ . Bankruptcy problems have been studied from a game theoretical viewpoint in O'Neill (1982). Given a bankruptcy problem  $(N, E, c)$ , the corresponding *bankruptcy game*,  $(N, v_{(N, E, c)})$ , is defined by

$$v_{(N, E, c)}(\mathcal{C}) = \max \left\{ 0, E - \sum_{i \in N \setminus \mathcal{C}} c_i \right\}$$

for every  $\mathcal{C} \subset N$ . In Curiel et al. (1987) it is shown that bankruptcy games are convex.

**Lemma 5.4** *Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem, let  $\mu_0$  be an initial matching function, and let  $\hat{\mu} \in \mathcal{M}^{max}$ . Then,  $(\mathcal{N} \cup \mathcal{D}, v^{\hat{\mu}})$  is a bankruptcy game.*

**Proof.** Define  $E(\hat{\mu}) = \sum_{i \in \mathcal{N}} d_i(\hat{\mu}) + \sum_{H \in \mathcal{D}} d_H(\hat{\mu})$ ,  $c_i = d_i^+(\hat{\mu})$  for every  $i \in \mathcal{N}$ , and  $c_H = d_H^+(\hat{\mu})$  for every  $H \in \mathcal{D}$ . We will prove that  $(\mathcal{N} \cup \mathcal{D}, E(\hat{\mu}), c)$  is a bankruptcy problem with the associated bankruptcy game  $v_{(\mathcal{N} \cup \mathcal{D}, E(\hat{\mu}), c)}$  equal to  $v^{\hat{\mu}}$ .

Note that  $(\mathcal{N} \cup \mathcal{D}, E(\hat{\mu}), c)$  is a bankruptcy problem since

$$\begin{aligned} E(\hat{\mu}) &= \sum_{i \in \mathcal{N}} d_i(\hat{\mu}) + \sum_{H \in \mathcal{D}} d_H(\hat{\mu}) = \sum_{i \in \mathcal{N}} d_i^+(\hat{\mu}) - \sum_{i \in \mathcal{N}} d_i^-(\hat{\mu}) + \sum_{H \in \mathcal{D}} d_H^+(\hat{\mu}) - \sum_{H \in \mathcal{D}} d_H^-(\hat{\mu}) \\ &\leq \sum_{i \in \mathcal{N}} d_i^+(\hat{\mu}) + \sum_{H \in \mathcal{D}} d_H^+(\hat{\mu}) = \sum_{i \in \mathcal{N}} c_i + \sum_{H \in \mathcal{D}} c_H \end{aligned}$$

by definition of  $d_i^-(\hat{\mu})$  and  $d_H^-(\hat{\mu})$ . Moreover, for  $\mathcal{C} \subset \mathcal{N} \cup \mathcal{D}$ , we have

$$\begin{aligned} v^{\hat{\mu}}(\mathcal{C}) &= \max \left\{ 0, \sum_{i \in \mathcal{C} \cap \mathcal{N}} d_i^+(\hat{\mu}) + \sum_{H \in \mathcal{C} \cap \mathcal{D}} d_H^+(\hat{\mu}) - \sum_{i \in \mathcal{N}} d_i^-(\hat{\mu}) - \sum_{H \in \mathcal{D}} d_H^-(\hat{\mu}) \right\} \\ &= \max \left\{ 0, \sum_{i \in \mathcal{N}} d_i^+(\hat{\mu}) + \sum_{H \in \mathcal{D}} d_H^+(\hat{\mu}) - \sum_{i \in \mathcal{N}} d_i^-(\hat{\mu}) - \sum_{H \in \mathcal{D}} d_H^-(\hat{\mu}) \right. \\ &\quad \left. - \sum_{i \in \mathcal{N} \setminus (\mathcal{C} \cap \mathcal{N})} d_i^+(\hat{\mu}) - \sum_{H \in \mathcal{D} \setminus (\mathcal{C} \cap \mathcal{D})} d_H^+(\hat{\mu}) \right\} \\ &= \max \left\{ 0, \sum_{i \in \mathcal{N}} d_i(\hat{\mu}) + \sum_{H \in \mathcal{D}} d_H(\hat{\mu}) - \sum_{i \in \mathcal{N} \setminus (\mathcal{C} \cap \mathcal{N})} d_i^+(\hat{\mu}) - \sum_{H \in \mathcal{D} \setminus (\mathcal{C} \cap \mathcal{D})} d_H^+(\hat{\mu}) \right\} \\ &= \max \left\{ 0, E(\hat{\mu}) - \sum_{i \in \mathcal{N} \setminus (\mathcal{C} \cap \mathcal{N})} c_i - \sum_{H \in \mathcal{D} \setminus (\mathcal{C} \cap \mathcal{D})} c_H \right\} \\ &= v_{(\mathcal{N} \cup \mathcal{D}, E(\hat{\mu}), c)}(\mathcal{C}). \end{aligned}$$

■

**Theorem 5.5** *Compensation matching games have a nonempty core.*

**Proof.** Let  $(\mathcal{N}, \mathcal{D}, q, r, \pi)$  be a matching problem, let  $\mu_0$  be an initial matching function, and let  $(\mathcal{N} \cup \mathcal{D}, v)$  be the corresponding compensation matching game. By Lemma 5.4, we know that  $v$  is the minimum of a finite number of bankruptcy games with the value of the grand coalition equal for all games. Clearly, this implies that  $\text{Core}(v^{\hat{\mu}}) \subseteq \text{Core}(v)$  for every  $\hat{\mu} \in \mathcal{M}^{\max}$ . Because bankruptcy games are convex, they have a nonempty core and, therefore,  $\text{Core}(v^{\hat{\mu}}) \neq \emptyset$  for every  $\hat{\mu} \in \mathcal{M}^{\max}$ , which implies  $\text{Core}(v) \neq \emptyset$ . ■

Theorem 5.5 tells us that we can find a compensation schedule such that a centralized restructuring resulting in a matching of maximum total reward of the organization will be ex-post stable, in the sense that there will be no subset of workers or departments who can by deviation obtain a higher reward. Still, the core of matching games may contain many possible allocations, thus, the question remains what allocation will be selected. The following example illustrates how such a selection could be conducted.

**Example 5.6** Reconsider the compensation matching game of Example 5.3. Recall that there are four matchings with maximum total reward:

$$\begin{aligned} \hat{\mu}_1: \quad & \hat{\mu}_1(1) = \hat{\mu}_1(2) = A, \quad \hat{\mu}_1(3) = \hat{\mu}_1(4) = B, \\ \hat{\mu}_2: \quad & \hat{\mu}_2(3) = \hat{\mu}_2(4) = A, \quad \hat{\mu}_2(1) = \hat{\mu}_2(2) = B, \\ \hat{\mu}_3: \quad & \hat{\mu}_3(1) = \hat{\mu}_3(3) = A, \quad \hat{\mu}_3(2) = \hat{\mu}_3(4) = B, \\ \hat{\mu}_4: \quad & \hat{\mu}_4(2) = \hat{\mu}_4(4) = A, \quad \hat{\mu}_4(1) = \hat{\mu}_4(3) = B. \end{aligned}$$

and that there is an extra reward of 12 when going from the initial matching to one of maximum total reward. Since  $v(\mathcal{C}) = \min \{v^{\hat{\mu}_1}(\mathcal{C}), v^{\hat{\mu}_2}(\mathcal{C}), v^{\hat{\mu}_3}(\mathcal{C}), v^{\hat{\mu}_4}(\mathcal{C})\}$  with  $v^{\hat{\mu}_1}(\mathcal{N} \cup \mathcal{D}) = v^{\hat{\mu}_2}(\mathcal{N} \cup \mathcal{D}) = v^{\hat{\mu}_3}(\mathcal{N} \cup \mathcal{D}) = v^{\hat{\mu}_4}(\mathcal{N} \cup \mathcal{D})$ , we have that  $\text{Core}(v^{\hat{\mu}_l}) \subset \text{Core}(v)$ . Moreover, since  $(\mathcal{N} \cup \mathcal{D}, v^{\hat{\mu}_l})$  is a bankruptcy game for every  $l \in \{1, 2, 3, 4\}$  by Lemma 5.4, the allocation given by any bankruptcy rule applied to the bankruptcy problem associated to  $\hat{\mu}_l$ ,  $l \in \{1, 2, 3, 4\}$ , is in the core of the compensation matching game, as well as any convex combination of those allocations. As an illustration, Table 3 gives the allocations obtained by applying the proportional rule (Prop), the adjusted proportional rule (AProp), the constrained equal awards rule (CEA), and the Talmud rule (Tal) to the bankruptcy problems as defined in the proof of Lemma 5.4. For a description of these rules see Thomson (2003).

If the implemented matching of maximal total reward is known a priori, one can use a specific bankruptcy rule, for example the proportional rule, in the corresponding bankruptcy problem associated to that selected matching. Otherwise, if a probability distribution over the possible acceptance of a matching of maximal total reward is known a priori, a possible allocation can be the expected allocation given by a specific bankruptcy rule. ♦

	$\hat{\mu}_1, \hat{\mu}_2 : (\mathcal{N}, 12, (6, 7, 3, 0))$	$\hat{\mu}_3, \hat{\mu}_4 : (\mathcal{N}, 12, (0, 6, 8, 5))$
Prop( $N, E, c$ )	$(4\frac{1}{2}, 5\frac{1}{4}, 2\frac{1}{4}, 0)$	$(0, 3\frac{15}{19}, 5\frac{1}{19}, 3\frac{3}{19})$
AProp( $N, E, c$ )	$(4\frac{6}{11}, 5\frac{6}{11}, 1\frac{10}{11}, 0)$	$(0, 3\frac{2}{3}, 5\frac{5}{18}, 3\frac{1}{18})$
CEA( $N, E, c$ )	$(4.5, 4.5, 3, 0)$	$(0, 4, 4, 4)$
Tal( $N, E, c$ )	$(4\frac{2}{3}, 5\frac{2}{3}, 1\frac{2}{3}, 0)$	$(0, 3\frac{2}{3}, 5\frac{2}{3}, 2\frac{2}{3})$

Table 3: Bankruptcy rules for compensation matching games in Example 5.6.

## References

- Alcalde, J. (1995), ‘Exchange-proofness or divorce-proofness? Stability in one-sided matching markets’, *Economic Design* **1**, 275–287.
- Bogomolnaia, A. and Jackson, M. (2002), ‘The stability of hedonic coalition structures’, *Games and Economic Behavior* **38**, 201–230.
- Chechlárová, K. (2002), ‘On the complexity of exchange-stable roommates’, *Discrete Applied Mathematics* **116**, 279–287.
- Chechlárová, K. and Manlove, D. (2005), ‘Exchange-stable marriage problem’, *Discrete Applied Mathematics* **152**, 109–122.
- Curiel, I. J., Maschler, M. and Tijs, S. (1987), ‘Bankruptcy games’, *Zeitschrift für Operations Research* **31**, A 143–A 159.
- Erdil, A. and Ergin, H. (2008), ‘What’s the matter with tie-breaking? Improving efficiency in school choice’, *American Economic Review* **98**(3), 669–689.
- Gale, D. and Shapley, L. S. (1962), ‘College admissions and the stability of marriage’, *American Mathematical Monthly* **69**, 9–15.
- O’Neill, B. (1982), ‘A problem of rights arbitration from the Talmud’, *Mathematical Social Sciences* **2**, 345–371.
- Shapley, L. S. (1971), ‘Cores of convex games’, *International Journal of Game Theory* **1**, 11–26.
- Thomson, W. (2003), ‘Axiomatic and game-theoretic analysis of bankruptcy and taxation problems: a survey’, *Mathematical Social Sciences* **45**, 249–297.