

Ágoston, Kolos Csaba; Biró, Péter; Szántó, Richárd

**Working Paper**

## Stable project allocation under distributional constraints

IEHAS Discussion Papers, No. MT-DP - 2017/33

**Provided in Cooperation with:**

Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences

*Suggested Citation:* Ágoston, Kolos Csaba; Biró, Péter; Szántó, Richárd (2017) : Stable project allocation under distributional constraints, IEHAS Discussion Papers, No. MT-DP - 2017/33, ISBN 978-615-5754-26-5, Hungarian Academy of Sciences, Institute of Economics, Budapest

This Version is available at:

<https://hdl.handle.net/10419/222010>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

---

MŰHELYTANULMÁNYOK

DISCUSSION PAPERS

**MT-DP – 2017/33**

**Stable project allocation  
under distributional constraints**

KOLOS CSABA ÁGOSTON - PÉTER BIRÓ - RICHÁRD SZÁNTÓ

Discussion papers  
MT-DP – 2017/33

Institute of Economics, Centre for Economic and Regional Studies,  
Hungarian Academy of Sciences

KTI/IE Discussion Papers are circulated to promote discussion and provoke comments.  
Any references to discussion papers should clearly state that the paper is preliminary.  
Materials published in this series may subject to further publication.

Stable project allocation under distributional constraints

Authors:

Kolos Csaba Ágoston  
Department of Operations Research and Actuarial Sciences,  
Corvinus University of Budapest  
Email: kolos.agoston@uni-corvinus.hu

Péter Biró  
senior research fellow  
Institute of Economics, Centre for Economic and Regional Studies,  
Hungarian Academy of Sciences  
and Department of Operations Research and Actuarial Sciences, Corvinus University  
Email: biro.peter@krtk.mta.hu

Richárd Szántó  
Department of Decision Sciences  
Corvinus University of Budapest  
Email: richard.szanto@uni-corvinus.hu

November 2017

ISBN 978-615-5754-26-5  
ISSN 1785 377X

# **Stable project allocation under distributional constraints**

Kolos Csaba Ágoston - Péter Biró - Richárd Szántó

## **Abstract**

In a two-sided matching market when agents on both sides have preferences the stability of the solution is typically the most important requirement. However, we may also face some distributional constraints with regard to the minimum number of assignees or the distribution of the assignees according to their types. These two kind of requirements can be challenging to reconcile in practice. In this paper we describe two real applications, a project allocation problem and a workshop assignment problem, both involving some distributional constraints. We used integer programming techniques to find reasonably good solutions with regard to the stability and the distributional constraints. Our approach can be useful in a variety of different applications, such as resident allocation with lower quotas, controlled school choice or college admissions with affirmative action.

**Keywords:** stable matching, two-sided markets, project allocation, linear programming, multi-criteria decision making

**JEL:** C61, C63, C78

## **Acknowledgements:**

A preliminary version of this paper has appeared in the proceedings of the 10th Japanese-Hungarian Symposium on Discrete Mathematics and its Applications, 2017.

Supported by the Hungarian Academy of Sciences under its Momentum Programme (LP2016-3/2016), by OTKA grant no. K108673.

# **Stabil projektallokáció disztribúciós feltételek mellett**

Ágoston Kolos Csaba - Biró Péter - Szántó Richárd

## **Kivonat**

Kétoldali párosítási piacoknál, amelyeknél mindkét oldal preferenciái számítanak, a megoldás stabilitása a legfontosabb feltétel. Lehetséges azonban, hogy emellett disztribúciós követelmények is megfogalmazódnak az allokált személyek minimális számára vagy az egyes típusú résztvevők elosztására vonatkozólag. Ezen kétfajta feltételt kihívás lehet összeegyeztetni a gyakorlatban. Ebben a tanulmányban két alkalmazást írunk le – egy projektallokációs problémát és egy workshop beosztást –, amelyekben disztribúciós feltételek álltak fent. A feladatok megoldására egészértékű programozási technikát használtunk a stabilitási és disztribúciós feltételek figyelembevételével. Megoldási módszerünk hasznos lehet más alkalmazásokban is, például rezidensek allokációjára alsó kvóták esetén, kontrollált iskolaválasztási programokban és központilag koordinált egyetemi felvételiben.

JEL: C61, C63, C78

**Tárgyszavak:** stabil párosítás, kétoldali piacok, projektallokáció, egészértékű programozás, többszemponú optimalizálás

# Stable project allocation under distributional constraints\*

Kolos Csaba Ágoston<sup>1</sup>, Péter Biró<sup>2†</sup> and Richárd Szántó<sup>3</sup>

<sup>1</sup> *Department of Operations Research and Actuarial Sciences,  
Corvinus University of Budapest, H-1093, Fővám tér 13-15., Budapest, Hungary  
Email: kolos.agoston@uni-corvinus.hu.*

<sup>2</sup> *Institute of Economics, Research Centre for Economic and Regional Studies,  
Hungarian Academy of Sciences, H-1112, Budaörsi út 45, Budapest, Hungary, and  
Department of Operations Research and Actuarial Sciences, Corvinus University of Budapest  
Email: peter.biro@krtk.mta.hu.*

<sup>3</sup> *Department of Decision Sciences  
Corvinus University of Budapest, H-1093, Fővám tér 13-15., Budapest, Hungary  
Email: richard.szanto@uni-corvinus.hu.*

**Abstract.** In a two-sided matching market when agents on both sides have preferences the stability of the solution is typically the most important requirement. However, we may also face some distributional constraints with regard to the minimum number of assignees or the distribution of the assignees according to their types. These two kind of requirements can be challenging to reconcile in practice. In this paper we describe two real applications, a project allocation problem and a workshop assignment problem, both involving some distributional constraints. We used integer programming techniques to find reasonably good solutions with regard to the stability and the distributional constraints. Our approach can be useful in a variety of different applications, such as resident allocation with lower quotas, controlled school choice or college admissions with affirmative action.

**Keywords:** stable matching, two-sided markets, project allocation, linear programming, multi-criteria decision making

## 1 Introduction

Centralised matching scheme has been used since 1952 in the US to allocate junior doctors to hospitals [36]. Later, the same technology has been used in school choice programs in large cities, such as New York [3] and Boston [4]. Similar schemes have been established

---

\*A preliminary version of this paper has appeared in the proceedings of the 10th Japanese-Hungarian Symposium on Discrete Mathematics and its Applications, 2017.

<sup>†</sup>Supported by the Hungarian Academy of Sciences under its Momentum Programme (LP2016-3/2016), by OTKA grant no. K108673.

in Europe for university admissions and school choice as well. For instance, in Hungary both the secondary school and the higher education admission schemes are organised nationwide, see [10] and [11], respectively. Furthermore, it can also be used to allocate courses to students under priorities [18]. In the above mentioned applications it is common that the preferences of the applicants and the rankings of the parties on the other side are collected by a central coordinator and a so-called stable allocation is computed based on the matching algorithm of Gale and Shapley [24]. Two-sided matching markets, and the above applications in particular, have been extensively studied in the last decades, see [39] and [32] for overviews from game theoretical and computational aspects, respectively.

In this paper we describe two recent applications at the Corvinus University of Budapest, where we used a similar method with some interesting caveats. In the first application we had to allocate students to projects in such a way that the number of students allocated to each project is between a lower and an upper quota, together with an additional requirement over the distribution of the foreign students. This is a natural requirement present in many applications, such as the Japanese resident allocation scheme [28]. In the second application we scheduled students to companies for solving case studies in a conference, and here again we faced some distributional constraints.

We decided to use integer programming techniques for solving both applications. We had at least three reasons for choosing this technique. The first is that with IP formulations we can easily encode those distributional requirements that the organisers requested, so this solution method is robust to accommodate special features. The second reason is that the computational problem became NP-hard as the companies submitted lists with ties. Using ties in the ranking was by our recommendation to the companies, because ties give us more flexibility when finding a stable solution under the distributional constraints. We describe this issue more in detail shortly. Finally, our third reason for choosing IP techniques was that it facilitates multi-objective optimisation, e.g. finding a most-stable solution if a stable solution does not exist under the strict distributional constraints.

The usage of integer programming techniques for solving two-sided stable matching problems is very rare in the applications, and the theoretical studies on this topic have only started very recently. The reason is that the problems are relatively large in most applications, and the Gale-Shapley type heuristics are usually able to find stable solutions, even in potentially challenging cases. A classical example is the resident allocation problem with couples, which has been present in the US application for decades, and it is still solved by the Roth-Peranson heuristic [38]. The underlying matching problem is NP-hard [35], but heuristic solutions are quite successful in practice, see also [12] on the Scottish application. However, integer programming and constraints programming techniques have been developed very recently and they turned out to be powerful enough to solve large random instances [13], [16] and [19]. Similarly encouraging results have been obtained for some special college admission problems, which are present in the Hungarian higher education system. These special features also makes the problem NP-hard in general, but at least one of these challenging features, turned out to be solvable even in a real data involving more than 150,000 applicants [6]. Finally, the last paper that we highlight with regard to this topic deals with the problem of finding stable solutions in the presence of ties [31]. However, we are not aware of any papers that would study IP techniques for the problem of distributional constraints.

Distributional constraints are present in many two-sided matching markets. In the

Japanese resident allocation the government wants to ensure that the doctors are evenly distributed across the country, and to achieve this they imposed lower quotas on the number of doctors allocated in each region [28, 29, 25]. Distributional objectives can also appear in school choice programs, where the decision makers want to control the socio-ethnic distribution of the students [2, 15, 20, 21]. Furthermore, the same kind of requirements are implemented in college admission schemes with affirmative action [1] such as the Brazilian college admission system [7] and the admission scheme to Indian engineering schools [8].

When stable solution does not exist for the strict distributional constraints then we either need to relax stability or to adjust the distributional constraints. In this study we will consider the trade-off between these two goals, and develop some reasonable solution concepts. We use integer programming technique to solve the problems arising from the two real applications and also for the simulations.

## 2 Definitions and preliminaries

Many-to-one stable matching markets have been defined in many context in the literature. In the classical college admissions problem by Gale and Shapley [24] the students are matched to colleges. In the computer science literature this problem setting is typically called Hospital / Residents problem (HR), due to the National Resident Matching Program (NRMP) and other related applications. In our paper we will refer the two sets as *applicants*  $A = \{a_1, \dots, a_n\}$  and *companies*  $\{C = c_1, \dots, c_m\}$ . Let  $u_j$  denote the upper quota of company  $c_j$ .

Regarding the preferences, we assume that the applicants provide strict rankings over the companies, but the companies may have ties in their rankings. This model is sometimes referred to as Hospital / Residents problem with Ties (HRT) in the computer science literature, see e.g. [32]. In our context, let  $r_{ij}$  denote the rank of company  $c_j$  in  $a_i$ 's preference list, meaning that applicant  $a_i$  prefers  $c_j$  to  $c_k$  if and only if  $r_{ij} < r_{ik}$ . Let  $s_{ij}$  be an integer representing the score of  $a_i$  by company  $c_j$ , meaning that  $a_i$  is preferred over  $a_k$  by company  $c_j$  if  $s_{ij} > s_{kj}$ . Note that here two applicants may have the same score at a company, so  $s_{ij} = s_{kj}$  is possible. Let  $\bar{s}$  denote the maximum possible score at any company and let  $E$  be the set of applications. A *matching* is a subset of applications, where each applicant is assigned to at most one company and the number of assignees at each company is less than or equal to the upper quota. A matching is *complete* if every student is allocated. A matching is said to be *stable* if for any applicant-company pair not included in the matching either the applicant is matched to a more preferred company or the company filled its upper quota with applicants of the same or higher scores.

In the classical college admission problem, that we refer to as HR, a stable solution is guaranteed to exist, and the two-versions of the Gale-Shapley algorithm [24] find either a student-optimal or a college optimal solution, respectively. Furthermore, this algorithm can be implemented to run in linear time in the number of applications. Moreover, the student-proposing variant was also proved to be strategyproof for the students [36], which means that no student can ever get a better partner by submitting false preferences. Finally, the so-called Rural Hospitals Theorem [37] states that the same students are matched in every stable solution, the number of assignees does not vary across stable



matchings for any college, and for the less popular colleges where the upper quota is not filled the set of assigned students is fixed.

When extending the classical college admission problem with the possibility of having ties in the colleges' rankings, that we referred to as an HRT instance, the existence of a stable solution is still guaranteed, since we can break the ties arbitrarily, and a stable solution for the strict preferences is also stable for the original ones. However, now the set of matched students and the size of the stable matchings can vary. Take just the following simple example: we have two applicants,  $a_1$  and  $a_2$  first applying to college  $c_1$  with the same score and applicant  $a_2$  also applies to college  $c_2$  as her second choice. Here, if we break the tie at  $c_1$  in favour of  $a_1$  then we get the matching  $a_1c_1, a_2c_2$ , whilst if we break the tie in favour of  $a_2$  then the resulting stable matching is  $a_2c_1$  (thus  $a_1$  is unmatched). The problem of finding a maximum size stable matching turned out to be NP-hard [33], and has been studied extensively in the computer science literature, see e.g. [32]. Note that when the objective of an application is to find a maximum size stable matchings, such as the Scottish resident allocation scheme [27], then the mechanism is not strategyproof. To see this, we just have to reconsider the above example, and assume that originally  $a_1$  also found  $c_2$  acceptable and would ranked it second, just like  $a_2$ . By removing  $c_2$  from her list,  $a_1$  is now guaranteed to get  $c_1$  is the maximum size stable solution, however, for the original true preferences  $a_2$  would have an equal chance to get her first choice  $c_1$ .

## 2.1 Introduction of lower quotas

In our first application the organisers of the project allocations wanted to ensure a minimum number of students for each company. Similar requirements have been imposed for the Japanese regions with regard to the number of residents allocated there. In our model, we introduce a lower quota  $l_j$  for each company  $c_j$  and we require that in a feasible matching the number of assignees at any company is between the lower and upper quotas. Stability is defined as before. We refer to the setting with strict preferences as Hospitals / Residents problem with Lower quotas (HRL) and the case with ties is referred to as Hospitals / Residents problem with Ties and Lower Quotas (HRTL).

Regarding HRL, the Rural Hospitals Theorem implies that the existence of a stable matching that obeys both the lower and upper quotas can be decided efficiently. This is because we just find one stable matching by considering the upper quotas only, and if the lower quotas are violated then there exists no stable solution under these distributional constraints. This problem can be still solved efficiently when the sets of companies have common lower and upper quotas in a laminar system, see [23].

However, the problem of deciding the existence of a stable matching for HRTL is NP-hard. To see this, we just have to remark that the problem of finding a complete stable matching for HRT with unit quotas is also NP-hard [33], so if we require both lower and upper quotas to be equal to one for all companies then the two problems are equivalent. Furthermore, no mechanism that finds a stable matching whenever there exists one can be strategyproof.

## 2.2 Adding types and distributional constraints

In our first application, the organisers want to distribute the foreign students across the projects almost equally. In our second application, there are target numbers for the

total number of Hungarian, European and other participants and there are also specific lower quotas for Hungarian students by some companies. These applications motivate our problems with applicant types and distributional constraints.

Let  $\mathcal{T} = \{T^1, \dots, T^p\}$  be the set of types, where  $t(a_i)$  denotes the type of applicant  $a_i$ . For a company  $c_j$ , let  $l_j^k$  and  $u_j^k$  denote the lower and upper quota for the number of assignees of type  $T^k$ . Furthermore, we may also set lower and upper quotas for any type of applicants for a set of companies. In particular, we denote the lower and upper quotas for the total number of applicants of type  $T^k$  assigned in the matching by  $L^k$  and  $U^k$ , respectively. The set of feasibility constraints for the matching is now extended with these lower and upper quotas. Yet, the original stability condition, which does not consider the types of the applicants, remains the same.

### 3 Solution concepts and integer programming formulations

In all of our formulations we use binary variables  $x_{ij} \in \{0, 1\}$  for each application coming from applicant  $a_i$  to company  $c_j$ . This can be seen as a characteristic function of the matching, where  $x_{ij} = 1$  corresponds to the case when  $a_i$  is assigned to  $c_j$ .

When describing the integer formulations, first we keep the stability condition fixed while we implement the set of distributional constraints. Then we investigate the ways one can relax stability or find most-stable solutions under the distributional constraints.

#### 3.1 Finding stable solutions under distributional constraints

In this subsection we gradually add constraints to the model by keeping the classical stability condition.

##### Classical HR instance

First we describe the basic IP formulation for HR described in [9]. The feasibility of a matching can be ensured with the following two sets of constraints.

$$\sum_{j:(a_i, c_j) \in E} x_{ij} \leq 1 \text{ for each } a_i \in A \quad (1)$$

$$\sum_{i:(a_i, c_j) \in E} x_{ij} \leq u_j \text{ for each } c_j \in C \quad (2)$$

Note that (1) implies that no applicant can be assigned to more than one company, and (2) implies that the upper quotas of the companies are respected.<sup>1</sup>

To enforce the stability of a feasible matching we can use the following constraint.

$$\left( \sum_{k:r_{ik} \leq r_{ij}} x_{ik} \right) \cdot u_j + \sum_{h:(a_h, c_j) \in E, s_{hj} > s_{ij}} x_{hj} \geq u_j \text{ for each } (a_i, c_j) \in E \quad (3)$$

---

<sup>1</sup>These conditions are standard for the assignment problem as well, see a survey on this problem and its variants [34] and an interesting application on marriage markets [17].

Note that for each  $(a_i, c_j) \in E$ , if  $a_i$  is matched to  $c_j$  or to a more preferred company then the first term provides the satisfaction of the inequality. Otherwise, when the first term is zero, then the second term is greater than or equal to the right hand side if and only if the places at  $c_j$  are filled with applicants with higher scores.

Among the stable solutions we can choose the applicant-optimal one by minimising the following objective function.

$$\sum_{(a_i, c_j) \in E} r_{ij} \cdot x_{ij}$$

### Modification for HRT

When the companies can express ties the following modified stability constraints, together with the feasibility constraints (1) and (2), lead to stable matchings. Note that here the only difference between this and the previous constraint is that the strict inequality  $s_{hj} > s_{ij}$  became weak.

$$\left( \sum_{k: r_{ik} \leq r_{ij}} x_{ik} \right) \cdot u_j + \sum_{h: (a_h, c_j) \in E, s_{hj} \geq s_{ij}} x_{hj} \geq u_j \text{ for each } (a_i, c_j) \in E \quad (4)$$

### Extension with lower quotas

Here, we only add the lower quotas for every company.

$$\sum_{i: (a_i, c_j) \in E} x_{ij} \geq l_j \text{ for each } c_j \in C \quad (5)$$

### Adding distributional constraints

As additional constraints we require the number of assignees of a particular type to be between the lower and upper quotas for that type at a company.

$$\sum_{i: t(a_i)=T^k, (a_i, c_j) \in E} x_{ij} \leq u_j^k \text{ for each } c_j \in C \text{ and } T^k \in \mathcal{T} \quad (6)$$

$$\sum_{i: t(a_i)=T^k, (a_i, c_j) \in E} x_{ij} \geq l_j^k \text{ for each } c_j \in C \text{ and } T^k \in \mathcal{T} \quad (7)$$

We can also add similar constraints for set of companies, or for the overall number of different assignees at all companies. We describe the latter, as we will use it when solving our second application.

$$\sum_{i, j: t(a_i)=T^k, (a_i, c_j) \in E} x_{ij} \leq U^k \text{ for each } T^k \in \mathcal{T} \quad (8)$$

$$\sum_{i, j: t(a_i)=T^k, (a_i, c_j) \in E} x_{ij} \geq L^k \text{ for each } T^k \in \mathcal{T} \quad (9)$$

### 3.2 Relaxing stability

Adding additional constraints to the problem can cause the lack of a stable matching, even if we added some flexibility with the ties.

One way to find a most-stable solution is to introduce nonnegative deficiency variables,  $d_{ij}$  for each application and add them to the left side of the stability constraint (4). By minimising the sum of these deficiencies as a first objective we can obtain a solution which is close to be stable.

$$\left( \sum_{k:r_{ik} \leq r_{ij}} x_{ik} \right) \cdot u_j + \sum_{h:(a_h, c_j) \in E, s_{hj} \geq s_{ij}} x_{hj} + d_{ij} \geq u_j \text{ for each } (a_i, c_j) \in E \quad (10)$$

Note that here, if a pair  $(a_i, c_j)$  is blocking for the assignment then we need to add more compensation  $d_{ij}$  if the number of assignees at  $c_j$  that the company prefers to  $a_i$  is large. This approach can be reasonable if we want to avoid the refusal of a very good candidate at a company. We call this solution as *matching with minimum deficiency*.

Alternatively, if we just want to minimise the number of blocking pairs then we can set  $d_{ij}$  to be binary and minimise the sum of these variables under the following modified constraints.

$$\left( \sum_{k:r_{ik} \leq r_{ij}} x_{ik} \right) \cdot u_j + \sum_{h:(a_h, c_j) \in E, s_{hj} \geq s_{ij}} x_{hj} + d_{ij} \cdot u_j \geq u_j \text{ for each } (a_i, c_j) \in E \quad (11)$$

Here, every blocking pair should be compensated by the same amount, so the number of blocking pairs is minimised. Note that this concept has already been studied in the literature for various models under the name of *almost stable matchings*, see e.g. [16].

### 3.3 Adjusting upper capacities, envy-free matchings

A different way of enforcing the lower quotas is to relax stability by artificially decreasing the capacities of the companies. This was also the solution in the resident allocation scheme in Japan [28], where the government introduced artificial upper quotas for each of the hospitals, so that in each region the sum of these artificial upper bounds summed up to the target capacity for that region. In the case of our motivating example of project allocation, one simple way of achieving the lower quotas was by reducing the upper quotas at every company.

In this solution what we essentially get is a so-called *envy-free matching*, studied in [42], [41] and [5]. The matching is stable with respect to the artificial upper quotas, which means that the only blocking pairs that may occur with regard to the original upper quotas are due to the empty slots created by the difference between the original and the artificial quotas, that we call *open-slot blockings*.

However, one may not want to reduce the upper quotas of the companies in the same way, perhaps some more popular companies should be allowed to have more students than the less popular ones. Furthermore, maybe the decision on which upper quotas should be reduced should be made depending on their effect of satisfying the lower quotas (or other

requirements). Thus, we may not want to set the artificial upper quotas in advance, but keep them as variables, by ensuring envy-freeness in a different way. One alternative way of enforcing envy-freeness is by the following set of constraints.

$$\sum_{k:r_{ik} \leq r_{ij}} x_{ik} \geq x_{hj} \quad \forall (a_i, c_j), (a_h, c_j) \in E, s_{ij} > s_{hj} \quad (12)$$

Constraints (12) will ensure envy-freeness, by making sure that if applicant  $a_h$  is assigned to company  $c_j$  and applicant  $a_i$  has higher score than  $a_h$  at  $c_j$  then  $a_i$  must be assigned to  $c_j$  or to a more preferred company.

### 3.4 Within-type priorities

So far we have only considered different approaches of relaxing stability or enlarging the set of feasible solutions in order to satisfy the distributional constraints. In this subsection we study alternative solution concepts and methods for the case when the distributional constraints are type-dependent. This is the case also in our motivating application, where special requirements are set for the foreign students assigned to the companies.

When the number of students of a type does not achieve the minimum required at a place then there are two well-known approaches. For instance in a school choice scenario, where the ratio of an socio-ethnic group should be improved (see e.g. [2]) then one possible affirmative action is to increase the scores of that group of students as much as needed. The other usual solution is to set some reserved seats to those students (see e.g. [7]).

In our project allocation application our requirement is to have at least one foreign student assigned to every company. If in a stable solution this condition would be violated for a company then we can try to enforce the admission of a foreign student by increasing the scores of the foreign students at this company. We call such a solution as *stable matching with type-specific scores*, where the classical stability condition is required for the adjusted scores. The second approach is to devote one place at each company to foreign students. For this one seat the foreign students will have higher priority than the locals irrespective of their scores, but for the rest of the spaces the usual score-based rankings apply. We call this concept as *stable matching with reserved seats for types*. Note that neither of these two concepts can always ensure that we get at least one foreign student at each company, since they may all have high scores and they may all dislike a particular company. However, this situation changes if we also allow to decrease the scores of a group of students. We will describe this case after discussing the third approach.

Finally, as a third approach, we can also extend the concept of envy-free matchings for types. We do not require any stability with regard to students of different types, but we do require envy-freeness for students of the same type. Thus the so-called *within-type envy-free matchings* will be those who satisfy the following set of constraints.<sup>2</sup>

$$\sum_{k:r_{ik} \leq r_{ij}} x_{ik} \geq x_{hj} \quad \forall (a_i, c_j), (a_h, c_j) \in E, s_{ij} > s_{hj}, t(a_i) = t(a_h) = T^k, T^k \in \mathcal{T} \quad (13)$$

That is, if  $a_i$  and  $a_h$  have the same type and  $a_h$  is assigned to  $c_j$  then the higher ranked  $a_i$  must also be assigned to  $c_j$  or to a more preferred company. Note that with

---

<sup>2</sup>This solution concept was called *within-type  $\succ$ -compatibility* by Echenique and Yenmez [20].

this modification we extend the set of feasible solutions compared to the set of envy-free matchings. Another important observation that is motivated by our project allocation problem is that under some realistic assumptions a within-type envy-free matching always exists, that we will show in the following theorem.

**Theorem 1** *Suppose that all the companies are acceptable to every student and that the sum of the lower quotas with regard to each type is less than equal to the number of students of that type, and the sum of the lower quotas across types for a company is less than or equal to the upper quota of that company, then a complete within-type envy-free matching always exists and can be found efficiently.*

*Proof:* We construct a within-type envy-free matching separately for each type and then we merge them. When considering a particular type  $T^k$ , we set artificial upper quotas at the companies to be equal to the type-specific lower quotas (i.e.  $l_j^k$  for company  $c_j$ ) and we find a stable matching  $M_k$  for this type. This stable matching must exist, since we assumed that all the companies are acceptable to every student and the number of students in every type is at least as much as the sum of the lower quotas for that type. We create matching  $M$  by merging the stable matchings for the types, i.e.  $M = M_1 \cup M_2 \cup \dots \cup M_p$ . Note that no upper quota is violated in  $M$ , since we assumed that the sum of the lower quotas across types for any company  $c_j$  is less than equal to the upper quota of  $c_j$ . By the stability of  $M_k$  for every type  $T^k$  it follows that matching  $M$  is within-type envy-free. If there is still a company  $c_j$ , where the overall lower quota ( $l_j$ ) is not yet met, then we increase an artificial upper quota for some at  $c_j$  so that there is still some unmatched applicants of this type. Since the total number of applicants is greater or equal to the sum of the lower quotas, we have to achieve the lower quotas at all companies in this way. Finally, if there are still some unmatched applicants then we increase some artificial upper quotas for their types one-by-one, by making sure that we never exceed any overall upper quota. At the end of this iterative process we must reach a complete within-type envy-free matching. ■

We note that there is a closely related solution concept introduced by Yokoi [41] which results in a within-type envy-free matching when restricted to our model, that we describe in details below. The model studied in that paper is the more general so-called classified stable matching problem where each student can have several types (e.g. gender, field of study, nationality) and the lower and upper quotas are set for every type. When putting their more general model in our context a student  $a_i$  has justified envy towards another student  $a_k$  at company  $c_j$  if  $a_k$  is assigned to  $c_j$ ,  $a_i$  prefers  $c_j$  to her assignment,  $c_j$  ranks  $a_i$  higher than  $a_k$ , and no lower and upper quota is violated for any type when replacing  $a_k$  with  $a_i$  at  $c_j$ . It is easy to see that under the assumptions of Theorem 1 an envy-free matching always exists as defined by Yokoi and such a solution is a within-type envy-free matching according to our definitions. Finally we remark that this model of Yokoi is originated from the classified stable matching problem introduced in [26], and further generalised in [23] and [40]. A common feature of these papers that the laminar nature of the set requirements makes the problem polynomial time solvable. A closely related model was studied in [22] without the laminar assumption, where the problem was proved to be NP-hard and was solved by integer programming techniques.

Let us abbreviate a *complete within-type envy-free matching* as CWTEFM. Now, we

will compare this concept of CWTEFM with stable matchings with type-specific scores and observe that they are essentially the same.

**Theorem 2** *Under the assumptions of Theorem 1 a complete matching is within-type envy-free if and only if it is stable with type-specific scores.*

*Proof:* Suppose first that  $M$  is a complete stable matching with type-specific scores, we will see that  $M$  is also within-type envy-free by definition. Suppose for a contradiction that there is a student  $a_i$  who has justified envy against student  $a_h$  of the same type at company  $c_j$ , i.e.  $a_h$  is assigned to  $c_j$  whilst  $a_i$  has higher score at  $c_j$  than  $a_h$  and  $a_i$  is assigned to a less preferred company. This would mean that the pair  $\{a_i, c_j\}$  is blocking for the adjusted scores, since both students get the same adjustment at  $c_j$ , contradicting with the stability of  $M$ .

Suppose now that  $M$  is a CWTEFM. Let us adjust the scores of the students according to their types at each company such that the weakest students admitted have the same scores across types. Matching  $M$  is stable with regard to the adjusted scores, because if a student  $a_i$  is not admitted to a company  $c_j$  and any better place of her preference that it must be the case that her score at  $c_j$  was less than or equal to the score of the weakest assigned student of the same type at  $c_j$ , which means that the adjusted score of  $a_i$  at  $c_j$  is less than or equal to the adjusted score of every assigned student at  $c_j$ . ■

Instead of using the above described processes of setting type-specific artificial upper quotas or making adjustments for the scores of different types, we can also get a CWTEFM directly by an IP formulation. We shall simply use the feasibility and distributional constraints together with (13) and with an objective function maximising the number of students assigned. This approach is not just more robust than the above described two heuristics, but it has also the advantage that we can enforce additional optimality or fairness criteria. As an additional fairness criterion we may aim to minimise the envy across types. We can achieve this by adding deficiency variables to the left hand side of constraints (12) for students of different types, as described in (14) below, and then minimising the sum of the deficiencies. We refer to this solution as Min#E-CWTEFM, that is *complete within-type envy-free matching with minimum number of envy across types*.

$$\sum_{k:r_{ik} \leq r_{ij}} x_{ik} + d_{ih}^j \geq x_{hj} \quad \forall (a_i, c_j), (a_h, c_j) \in E, t(a_i) \neq t(a_h) \quad (14)$$

However, we may find an envy more justified, if the score difference between the two applicants involved is higher. Thus, by taking the score differences as the intensities of the envies, we can also aim to find a refined solution where the total intensities of the envies is minimised, by using the following objective function:

$$\sum (s_{ij} - s_{hj}) \cdot d_{ih}^j.$$

We call the corresponding solution *complete within-type envy-free matching with minimum envy intensities across types*, abbreviated as MinEI-CWTEFM.

If the solution is still not unique then we can further refine it, by considering two additional objectives. Regarding the welfare of the students, we may want to minimise

the total rank of the students, leading to a Pareto-optimal assignment for them under the constraints. We denote these solutions as MinRank-Min#E-CWTEFM and MinRank-MinEI-CWTEFM, depending whether we minimised the number of envies or the envy intensities in the previous round. Finally, an alternative objective can be to minimise the number of blocking pairs due to open slots. This can be achieved by adding binary deficiency variables to the first term of the left side of the stability constraints, as follows.

$$\left( \sum_{k:r_{ik} \leq r_{ij}} x_{ik} + d_{ij} \right) \cdot u_j + \sum_{h:(a_h, c_j) \in E} x_{hj} \geq u_j \text{ for each } (a_i, c_j) \in E \quad (15)$$

We can then minimise the sum of these deficiency variables and find a matching within the restricted solution set that minimises the number of open-slot blockings. We denote these solutions as MinOSB-Min#E-CWTEFM and MinOSB-MinEI-CWTEFM, depending whether we minimised the number of envies or the envy intensities.

## 4 First application: CEMS project allocation

CEMS Alliance is a global co-operation of leading business schools, multinational corporations and social partners in higher education domain. These entities run together the CEMS Master in International Management (MIM) one-year graduate program that is accessible for graduate students of the partner institutions in 29 countries in five continents. During the one-year-program students spend one semester at their home institution and one semester at another partner institution somewhere abroad, and they always learn in an international environment. CEMS MIM has been ranked as a leading master program by Financial Times in recent years.

Within the framework of the MIM program each student must carry out a business project during the Spring semester accounting for 15ECTS credits (that is half of the workload of the entire semester). The consultancy-like projects are designed as real life learning experience. Business projects are done in small groups of 3-6 students in which ideally at least one student comes from a foreign school, hence business project teams are culturally diverse. Business projects are offered and supervised by the corporate partners throughout the semester and they usually last for three months.

Students learn about the business projects during a kickoff event at the beginning of the semester from company representatives and they also receive written descriptions of the projects. After the kickoff event corporate partners evaluate all students according to their CV-s, and students also rank the business projects in the same time. The school assigns students to the individual projects based on these evaluations and rankings.

At Corvinus University of Budapest the authors of this paper have been given the task of redesigning the allocation mechanism in 2016. In previous years the mechanism was a simple immediate acceptance mechanism (also known as Boston mechanism [4]), where the students submitted their CV-s to their first choice companies, the companies evaluated the candidates and then they accepted the best candidates up to their quotas and rejected the rest. The rejected students then submitted their CV-s to further companies, but those companies which have already filled their positions did not accept more applications. This mechanism was heavily criticized in the literature on school choice due to its unfairness



and also because this mechanism is highly manipulable, therefore in many cities it has been replaced by other algorithms, mainly by the deferred acceptance (or Gale-Shapley) algorithm, see e.g. [4].

#### 4.1 Solution plan

In 2016 there were 25 students, including 20 local and 5 foreign students, and 5 companies. The initial upper quotas were set to 6 and the lower quotas were set to 4 at all companies. The programme coordinator decided to set an upper quota of 2 for the foreign students at each company to enforce diversity. In 2017 there was a slight change in the distributional criteria, the number of students allocated to each company was set to be between 3 and 6 and at least one foreign student was required to be allocated to every company.

Our first solution plan was to ask the students to rank all the companies in a strict order and to ask the companies to evaluate all the CV-s and rank the students weakly by giving them scores between 1 and  $10^3$ . Our intention with allowing ties was to enlarge the set of stable solutions, even though we understand that this fairness concept is a bit weaker, since we may accept a student and reject another one with the same score. Allowing ties also makes the problem NP-hard already with lower quotas, as we described in the introduction. Yet, if the ties were not allowed then the set of stable (and envy-free) solutions would be much smaller and thus it would be harder to satisfy the distributional constraints.

We remark that the conditions of Theorem 1 are satisfied for both 2016 and 2017, since all the students have to rank (and accept) all the companies and in 2017 we were required to have at least one foreign student at each company, where the number of foreign students was more than the number of companies. Therefore a complete within-type envy-free matching always existed. Within this set of solutions we decided to minimise the number of envies across types and their intensities as the primal objectives. As secondary objectives we tried to minimise the total rank and the number of open-slot blockings.

Finally, since in both years it was possible to decrease the upper quotas at all companies by one (and set them to 5 instead of 6), we also examined these solutions. This was reasonable as allocating very different numbers of students to the companies seemed problematic, especially if some of the most popular companies was forced not to fill its quota, while less popular companies did.

#### 4.2 Results in 2016

The most important results of the 2016 matching run are collected in Table 1.

In 2016 the upper bound of two for the foreign students were always satisfied without considering it, so we leave out this question from the discussion and we focus only on common lower quotas. We were not able to find a stable solution for the original quotas of 4-6, since one of the companies (number 2) was very unpopular and the highest number of students that we could match there in a stable solution was 2, this is Solution 2 in Table 1. (For the record, we also checked which would be the minimum rank solution among the stable ones, that is Solution 1.) Therefore we decreased the upper quotas of all companies to 5, except

---

<sup>3</sup>Most companies gave only integer scores, but some submitted half-integer scores as well, so ties indeed occurred.

2016	profiles all/foreign					total rank
Solution 1: MinRank-Stable $u_i = 6$	6	1	6	6	6	34
Solution 2: MinRank-Stable $l_i = 2, u_i = 6$	6	2	6	6	5	35
Solution 3: MinRank-Stable $u_1 = 6, u_i = 5 (i = 2..5)$	6	4	5	5	5	40
Solution 4: MinRank-Stable $u_i = 5$	5	5	5	5	5	41
Solution 5: MinRank-EF $l_i = 4, u_i = 6$	5	4	6	6	4	38
Solution 6: MinOSB-EF $l_i = 4, u_i = 6$	6	4	6	5	4	39

Table 1: The results of the 2016 matching run with the number of all and foreign students assigned to the companies and the total rank of the students.

the most popular company (number 1) and found a stable matching with minimum total rank (Solution 3). Note that this matching is envy-free for the original quotas. Finally we considered the possibility of decreasing all the upper quotas to 5, as described in Solution 4. From the latter two solutions the decision maker decided to choose Solution 4, since it was not substantially different from Solution 3 and for the companies it seemed to be easier to communicate the common decrease of upper quotas, compared to the case when only one company has a larger number of students.

Recently, after carefully investigating the solution concepts described in this paper, we did another check on the possible results and computed Solutions 5 and 6. Solution 5 is an envy-free solution where the total rank is minimised. It was interesting to observe that the most popular company (number 1) does not fill its upper quota, leading to many open-slot blockings at that company. Solution 6 is also envy-free, but here the open-slot blockings are minimised, but this resulted in a small decrease in the total rank.

### 4.3 Results in 2017

The results of the 2017 matching run are summarised in Table 2. In 2017 the number of students was 40 among which 13 were from abroad and the number of companies was 8. Due to the higher proportion of foreign students, the organisers decided to require the allocation of at least one foreign student to each company. The initial call suggested groups of sizes between 3 and 6, but in this year also we investigated the solutions when every upper quota was decreased to 5. In the latter case the lower quotas for the foreign students were not automatically satisfied, so we found within-type envy-free solutions and then as a first objective we either minimised the number of envies across types or we minimised the intensities of the envies. As a secondary objective we tried to minimise the total rank (there was no open-slots blocking when the upper quotas were commonly set to 5).

2017	profiles all/foreign								total rank
Solution 1: MinRank-EF $l_i = 3, u_i = 6$	6	3	6	6	6	3	6	4	66
Solution 2: MinRank-Min#E-CWTEFM $u_i = 5, \text{wEF}, \text{min}$	5	5	5	5	5	5	5	5	85
Solution 3: MinRank-MinEI-CWTEFM $u_i = 5$	5	5	5	5	5	5	5	5	105

Table 2: The results of the 2017 matching run with the number of all and foreign students assigned to the companies and the total rank of the students.

Solution 1 is envy-free, and the total rank is minimised. As intuitively expected, the two least popular companies have only three students allocated each and a medium popular company has four students, whilst the popular companies receive six students. Solutions 2 and 3 are both within-type envy free for upper quotas 5. Solution 2 minimises the number of envies as the first objective and then the total rank. Solution 3 minimises the intensities of the envies and then the total rank. (Note that we also computed the minimal envy solutions without requiring within-type envy freeness, and essentially we received the same two solutions.) It is interesting to know that only one justified envy was present in both Solutions 2 and 3, and the intensity of this envy was 1 in Solution 2 and  $\frac{1}{2}$  in Solution 3. However, these two solutions were rather different, and Solution 2 had much smaller total rank. Thus Solution 2 was clearly found better than Solution 3 by the decision maker. When comparing the first two solutions, the decision maker selected Solution 2, due to the more balanced sizes of groups.

#### 4.4 Discussion, further questions

Here we discuss our findings and possible questions for the future.

**Importance of the distributional requirements.** We have considered our distributional constraints as hard bounds, the only relaxation we tested was the common decrease of the upper quotas. However, in many applications the distributional goals are softer, and thus may be violated. For instance, in school choice the exact proportionality with regard to ethnicity or gender may be too demanding and unnecessary to satisfy, these are rather just general aims. In such situations one may insist of the stability or the envy freeness of the solution and want to satisfy the distributional constraints as much as possible. Finally, the trade-off between fairness and distributional goals may be balanced by relaxing both requirements at the same time.

**Stability versus envy-freeness.** Leaving some slots empty to satisfy the distributional constraints is a natural way to relax stability. This is also used in the Japanese resident allocation programme, where artificial upper quotas have been set to the hospitals in order to satisfy the regional lower quotas [28]. However, the open-slot blocking can also be seen as unfair from both the students' and the companies' points of views, especially when a popular company has to give up an intern. Note also that the open-slot blockings are relative to the original quotas. In our application the decision maker ended up choosing solutions in both years where the upper quotas of the companies were commonly reduced

by one. These solutions admit a high number of open-slot blockings with regarding the original quotas, whilst if they are envy free for the original quotas then they are also stable (with no open-slot blockings) for the decreased quotas. Thus these chosen solutions can be seen more fair from the students' point of view, as they do not regret their rejections by a company with an open slot.

**Importance of within-type envy-freeness:** In our analyses we assumed that within-type envy-freeness is an important requirement that we obeyed in all solutions. Note that in the 2017 run we also tested the solution when this requirement was relaxed and we did not find a significant difference in the solutions. It is an interesting question how important this requirement is, and the answer can depend on the actual application. If the separation of the types is significant and there is a big difference between their performance (e.g. regarding the ethnicity in college admission) then within-type envy-freeness can be crucial.

**Minimising the number of justified envies or their intensities.** In our 2017 run we had a significant difference between our two recommended solutions based on minimising the number of justified envies and their intensities, respectively. In our case the former solution had much better total ranking for the students, but one can easily create an example where the opposite would happen. If the intensities of the blocking are minimised then this means that the average difference between the scores of the students who have envy towards one to another is small. This can be more acceptable than having large score differences. In fact, if the maximum score difference is not higher than one in our application, then we could say that this solution could be seen weakly stable if the scoring by the companies were less finer, say used score range 1-5 instead of the current range of 1-10.

**Strict versus weak rankings.** Using ties in the rankings of the companies was by our recommendation in order to enlarge the set of stable (or envy-free) matchings. However, in this case stability (and envy-freeness) is weaker, the rejection of a student by a company can be explained by the admittance of another student with equal score or higher. Thus, this can be seen unfair by the student rejected, therefore in many applications (e.g. school choice in New York, Boston and college admissions in Ireland and Turkey) the ties are broken by lotteries or by other random factors. Ties make the problem of satisfying lower quotas NP-hard, whilst this is a polynomial-time solvable problem for strict rankings, see e.g. [23]. Furthermore, the mechanism can become highly manipulable by the students for ties depending on the goals of the optimisation.

**Incentive issues.** A mechanism is strategy-proof for the students if neither of them can get a better match by submitting false preferences. This property holds for the student-proposing deferred-acceptance mechanism in the classical college admission model of Gale and Shapley (see e.g. [39]). Strategy-proofness can also be satisfied by modified variants of the deferred-acceptance mechanism for the case of lower quotas, as suggested also for the Japanese resident allocations [28, 29, 25]. However, if we allow ties and we consider goals such as rank-minimisation then our mechanism becomes manipulable. A simple manipulation strategy for a medium-strong student can be to put her top choice as first choice, but instead of putting her true second choice in the second slot she can put some companies which are not achievable for her in any stable solution. If there is another student with the very same score and very same preferences submitting her true preferences and there is only one place left at their most preferred company then the

rank-maximising algorithm with assign the manipulating student there, and the truth-telling student to the second company, since the alternative solution by exchanging the two students would result in higher total rank. Despite of this issue of manipulability, we believe that the expected gains of manipulations are negligible and their risks can be high, so in a Bayesian sense it is unlikely that a student could get a positive expected gain by manipulating. However, we admit that this hypothesis would be very hard to prove formally.

**Bound the length of preference lists.** In 2016 there were 25 students and 5 companies, in 2017 there were 40 students and 8 companies, so the screening costs of the companies have increased a lot. If this tendency will continue then the organisers of the programme may need to reconsider the requirement of providing full rankings. A reasonable solution in such situations is to have two rounds. In the first the student are required to rank a fixed number of companies, say five), and it is not guaranteed that all the students can be allocated to acceptable companies that they ranked. In the second round either no preferences are asked from the students or the organisers can elicit the preferences of the unmatched students over the companies with remaining positions. This is a standard technique also in school choice (e.g. in New York [3]), although here we would face new challenges to ensure the satisfaction of the distributional requirements.

## 5 Second application: Workshop assignment

After running the 2016 project allocation, we received very positive feedbacks from the students, and in fact two students approached us asking for a help in selecting and assigning conference participants to companies involved in a case study workshop within the conference.

The number of participants to be selected was 60, and they had to be assigned to three companies in a given proportion, the first company had to receive 16 students and at least 8 Hungarians, the second and the third companies had to receive 22 students each. There were 13 pre-selected students (the country leaders of the organisation) whose assignments were fixed in advance, so we only had to select and assign the 47 remaining slots.

The conference organisers also agreed on the proportion of the local, regional and other students to be selected. In particular, we had to select 25 Hungarian students from the 29 Hungarian applicants, further 12 regional students from the 15 regional applicants (outside Hungary) and 10 other students from the 19 other applicants (outside the region). Thus, we had overall exact quotas (i.e. equal lower and upper quotas) for each type of students, just as described in 8 and 9.

In order to satisfy these requirements we thought that we not only try to keep the solution within-type envy-free, which is also stable with type-specific scores as proved in Theorem 2, but we tried to keep the extra scores given to each type of students be the same across companies. We call this solution concept a *stable matching with equal type-specific scores*. With an iterative testing we could indeed find such a stable solution by adding 7 extra points to all Hungarian students, 3 extra points to all regional students, and zero to the other students, where the students had to rank all the three companies and the companies gave scores (1-10) on all the applicants.

It is an interesting question whether a stable matching with equal type-specific scores

always exists in our model, under the assumption that all pairs are acceptable. We state this as a conjecture below and prove it for two types.

**Conjecture 3** *When all the pairs are acceptable then a stable matching with equal type-specific scores always exists for exact quotas.*

To prove the conjecture for two types, we will use some well-known theorems listed below.

**Theorem 4** *[Well-known results on HR/HRT instances]*

- i) (Characterisation, see e.g. [27]) *A matching  $\mu$  is weakly stable for an instance  $I$  of HRT if and only if it is stable for an instance  $I'$  of HR that is obtained by some tie-breaking from  $I$ .*
- ii) (Rural hospitals [37]) *For an instance  $I$  of HR the set of allocated students and the number of seats filled at the companies are fixed across the stable matchings.*
- iii) (Vacancy chains [14]) *Suppose that  $I$  is an instance of HR. If  $I'$  is obtained from  $I$  by adding a new student  $a_i$  then the set of allocated students either 1) remains the same, 2) it is extended by  $a_i$ , or 3) it is extended by  $a_i$  and another student,  $a_j$  becomes unallocated. If  $I'$  is obtained from  $I$  by increasing the upper quota of a company then the set of allocated students either 1) remains the same, or 2) it is extended by one student.*

*Proof:*[of Conjecture 3 for two types] Suppose that we have a project allocation problem with two types of students,  $A_1 = \{a_1, a_2, \dots, a_{n_1}\}$  and  $A_2 = \{a_{n_1+1}, a_{n_1+2}, \dots, a_{n_1+n_2}\}$ , companies  $C = \{c_1, \dots, c_m\}$  and exact quotas  $L^1 = U^1$  and  $L^2 = U^2$ . Let  $U$  denote the total capacity, i.e.,  $U = \sum_{j=1..m} u_j$ . Without loss of generality we suppose that  $L^1 + L^2 = U$  with  $L^1 \leq n_1$  and  $L^2 \leq n_2$ , and  $m \leq n_1 + n_2$ , which implies that all the weakly stable matchings have size  $U$ , since we assume that every student-company pair is mutually acceptable. Let  $e$  denote the extra score given to students of the first type. Note that if  $e$  is a high number then all the first type students are admitted up to the total quota and if  $e$  is very small (negative) then all the second type student are admitted up to the total quota. However, the number of first type students admitted does not necessarily grow when we increase  $e$ .<sup>4</sup> Let  $I$  denote the original instance of HRT and let  $I_e$  be the instance of HRT obtained after adding point  $e$  to all students in  $A^1$ . For a matching  $M$  let  $|M|_1$  denote the number of first type student allocated and similarly, let  $|M|_2$  denote the number of second type students allocated in  $M$ . The goal is to find a suitable extra score  $e$  such that there exists a weakly stable matching  $M$  for  $I_e$  such that  $|M|_1 = L^1$  (which implies that  $|M|_2 = L^2$ ). In fact, we will construct a HR instance  $I'_e$ , obtained from  $I_e$

---

<sup>4</sup>The following example illustrates this. Let us have  $A^1 = \{a_1, a_2, a_3\}$  and  $A^2 = \{a_4, a_5\}$ , and  $C = \{c_1, c_2, c_3\}$  with quota 1 at each company. Suppose that every student prefers  $c_1$  to  $c_2$  and  $c_2$  to  $c_3$ . The students have the following scores:  $s_{1,1} = 5$ ,  $s_{1,2} = 7$ ,  $s_{1,3} = 1$ ,  $s_{2,1} = 1$ ,  $s_{2,2} = 1$ ,  $s_{2,3} = 3$ ,  $s_{3,1} = 1$ ,  $s_{3,2} = 1$ ,  $s_{3,3} = 1$ ,  $s_{4,1} = 6$ ,  $s_{4,2} = 1$ ,  $s_{4,3} = 6$ ,  $s_{5,1} = 2$ ,  $s_{5,2} = 6$ ,  $s_{5,3} = 2$ . When no extra score is added to  $A^1$  then the unique stable matching is  $M = \{a_1c_2, a_2c_3, a_4c_1\}$ , and when we increase the score of students in  $A^1$  then the unique stable matching is  $M' = \{a_1c_1, a_4c_3, a_5c_2\}$ , thus the number of first type students allocated decreases.

by tie-breaking, such that matching  $M$  is stable for  $I'_e$  with the required distributional property.

As we already noted, if  $e$  is a large negative number then  $|M|_1 = 0$  for any stable matching  $M$  in  $I_e$  and if  $e$  is a very large positive number then  $|M|_1 = \min\{n_1, m\}$ . For instance  $I_e$  of HRT, let  $I_e^{<1}$  denote the HR instance where all the ties are broken in favour of  $A^1$  students (and among the students of the same type we use an arbitrary tie-breaking, say, according to their indices). Similarly, let  $I_e^{<2}$  denote the HR instance, where we break all the ties in favour of  $A^2$  students. First, we have to observe that  $I_e^{<1}$  is the same as  $I_{e+1}^{<2}$  for any  $e$ . Note also that the number of allocated first type students (and their set) is fixed for any HR instance by Theorem 4/ii) across all stable matchings. Therefore there must exist a number  $e$  such that for any stable matching  $M_{e-1}$  for instance  $I_{e-1}^{<1}(= I_e^{<2})$  we have  $|M_{e-1}|_1 \leq L^1$ , and for any stable matching  $M_e$  for instance  $I_e^{<1}$  we have  $|M_e|_1 \geq L^1$ . We will show that there is an instance  $I'_e$ , obtained by tie-breaking from  $I_e$ , such that for every stable matching the number of allocated  $A^1$  students is exactly  $L^1$ .

We start from  $I_e^{<2}$  and we will gradually transform it into  $I_e^{<1}$  by giving higher priority in the tie-breaking to one  $A^1$  student in each step. Let  $I_e^0 = I_e^{<2}$ , and for each  $i \in [1..n_1]$  let  $I_e^i$  denote the HR instance where we favour the students  $\{a_1, \dots, a_i\}$  over students in  $A^2$ , who are favoured to student in  $\{a_{i+1}, \dots, a_{n_1}\}$ . What we will show is that if  $M$  is any stable matching for  $I_e^i$  and  $M'$  is any stable matching for  $I_e^{i+1}$  then  $|M|_1 - 1 \leq |M'| \leq |M|_1 + 1$ , so the number of  $A^1$  students allocated can either increase or decrease by at most one. This will imply that we must get an instance  $I'_e$ , where the number of first type students is exactly  $L^1$ .

To prove the above inequalities we have to consider two situations. First, let us assume that  $a_{i+1}$  is unmatched in  $M$  (and so in every stable matching for  $I_e^i$ ). Thus  $M$  would also be stable if we would remove  $a_{i+1}$  from  $I_e^i$ . Let us now put back  $a_{i+1}$ , but with higher priority, creating instance  $I_e^{i+1}$ . By Theorem 4/iii) either the number of  $A^1$  students remain the same or it increases by one. Suppose now that  $a_{i+1}$  is allocated in  $M$  to company  $c_j$ .  $M$  will remain stable if we remove both  $a_{i+1}$  and one seat at  $c_j$  from  $I_e^i$ , while the number of  $A^1$  students allocated in the reduced matching decreases by one. If we add back one seat at  $c_j$  and subsequently we also add back  $a_{i+1}$  with increased priority, creating instance  $I_e^{i+1}$ , then from Theorem 4/iii) we know that in each of these two steps the number of  $A^1$  students matched either remains the same or increases by one. So, in overall, the number of  $A^1$  students can either decrease by one, remain the same, or increase by one. This completes our proof.  $\blacksquare$

## 6 Conclusion

We investigated different solution concepts for stable matching problems with distributional constraints motivated by two real applications where we had to design the allocation mechanism. We chose integer programming as the solution technique which proved to be successful for these relatively small applications. We believe that our solution concepts and techniques could be considered in other applications as well, such as controlled school choice and university admission with affirmative action. As far as the participants are concerned, we have received very positive feedbacks from both the students and the companies, especially compared to the previous years. There are still plenty of interesting

questions to investigate mostly about the importance of different fairness criteria and the trade-off between fairness and the distributional requirements.

## References

- [1] Abdulkadiroglu A.. College admissions with affirmative action. *International Journal of Game Theory*, 33(4):535–549, 2005.
- [2] Abdulkadiroglu A., and Ehlers L. Controlled School Choice. *Working paper*, 2007.
- [3] Abdulkadiroğlu A., Pathak A., and Roth A.E. The New York City high school match. *American Economic Review, Papers and Proceedings*, 95(2):364–367, 2005.
- [4] Abdulkadiroğlu A., Pathak P.A., Roth A.E., and Sönmez T. The Boston public school match. *American Economic Review, Papers and Proceedings*, 95(2):368–371, 2005.
- [5] Ágoston K.Cs., and Biró P. Modelling Preference Ties and Equal Treatment Policy. *In Proceedings of ECMS 2017: 31st European Conference on Modelling and Simulation*, pages 516-522, 2017.
- [6] Ágoston K.Cs., Biró P., and McBride I. Integer programming methods for special college admissions problems. *Journal of Combinatorial Optimization*, 32(4):1371–1399, 2016.
- [7] Aygün O., and Bo I. College Admission with Multidimensional Reserves: The Brazilian Affirmative Action Case. *Working paper*, 2013.
- [8] Aygün O., and Turhan B. Dynamic reserves in matching markets: Theory and applications. *Working paper*, 2016.
- [9] Baïou M., and Balinski M. The stable admissions polytope. *Mathematical Programming*, 87(3), Ser. A:427–439, 2000.
- [10] Biró P. Matching Practices for Secondary Schools – Hungary. [matching-in-practice.eu](http://matching-in-practice.eu), accessed on 23 August 2014
- [11] Biró P. University admission practices - Hungary. [matching-in-practice.eu](http://matching-in-practice.eu), accessed on 23 August 2014
- [12] Biró P., Irving R.W., and Schlotter I. Stable matching with couples – an empirical study. *ACM Journal of Experimental Algorithmics*, 16, Article No.: 1.2, 2011.
- [13] Biró P., McBride I., and Manlove D.F. The Hospitals / Residents problem with Couples: Complexity and Integer Programming models. *In Proceedings of SEA 2014: the 13th International Symposium on Experimental Algorithms*, vol 8504 of LNCS, pp 10-21, Springer, 2014.
- [14] Blum J., Roth A.E., and Rothblum U.G. Vacancy chains and equilibration in senior-level labor markets. *ACM Journal of Experimental Algorithmics*, 76(2):362–411, 1997.



- [15] Bo I. Fair implementation of diversity in school choice. *Games and Economic Behavior*, 97:54–63, 2016.
- [16] McBride I., Manlove D.F., and Trimble J. "Almost stable" matchings in the Hospitals / Residents problem with Couples. *Constraints*, 22(1):50–72, 2016.
- [17] Cao N.V., Fragnière E., Gautier J.-A., Sapin M., and Widmer E.D. Optimizing the marriage market: An application of the linear assignment model. *European Journal of Operational Research*, 202(2):547–553, 2010.
- [18] Diebold F., and Bichler M. Matching with indifference: A comparison of algorithms in the context of course allocation. *European Journal of Operational Research*, 260(1):268–282, 2017.
- [19] Drummond J., Perrault A., and Bacchus F. SAT is an effective and complete method for solving stable matching problems with couples. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI-15)*, 2015.
- [20] Echenique F., and Yenmez M.B. How to control controlled school choice. *American Economic Review*, 105(8):2679–2694, 2015.
- [21] Ehlers L., Hafalir I.E., Yenmez M.B., and Yildirim M.A. School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory*, 153:648–683, 2014.
- [22] Firat M., Briskorn M., and Laugier A. A Branch-and-Price algorithm for stable workforce assignments with hierarchical skills. *European Journal of Operational Research*, 251:676–685, 2016.
- [23] Fleiner T., and Kamiyama N. A Matroid Approach to Stable Matchings with Lower Quotas. *Mathematics of Operations Research*, 41(2):734–744, 2016.
- [24] Gale D., and Shapley L.S. College Admissions and the Stability of Marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.
- [25] Goto M., Kojima F., Kurata R., Tamura A., and Yokoo M. Designing Matching Mechanisms Under General Distributional Constraints. *American Economic Journal: Microeconomics*, forthcoming, 2017.
- [26] Huang C.-C. Classified stable matching. In *Proceedings of SODA 2010: the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 1235–1253, 2010.
- [27] Irving R.W., and Manlove D.F. Approximation algorithms for hard variants of the stable marriage and hospitals/residents problems. *Journal of Combinatorial Optimization*, 16(3):279–292, 2008.
- [28] Kamada Y., and Kojima F. Efficient Matching Under Distributional Constraints: Theory and Applications. *American Economic Review*, 105(1):67–99, 2014.
- [29] Kamada Y., and Kojima F. Stability concepts in matching under distributional constraints. *Journal of Economic Theory*, 168:107–142, 2017.

- [30] Kojima F. School choice: Impossibilities for affirmative action. *Games and Economic Behavior*, 75(2):685–693, 2012.
- [31] Kwanashie A., and Manlove D.F. An Integer Programming approach to the Hospitals / Residents problem with Ties. In *Proceedings of OR 2013: the International Conference on Operations Research*, pages 263-269, Springer, 2014.
- [32] Manlove D.F. Algorithms of Matching Under Preferences. *World Scientific Publishing*, 2013.
- [33] Manlove D.F., Irving R.W., Iwama K., Miyazaki S., and Morita Y. Hard variants of stable marriage. *Theoretical Computer Science*, 276(1-2):261–279, 2002.
- [34] Pentico D.V. Assignment problems: A golden anniversary survey. *European Journal of Operational Research*, 176:774–793, 2007.
- [35] Ronn E. NP-complete stable matching problems. *Journal of Algorithms*, 11:285–304, 1990.
- [36] Roth A.E. The evolution of the labor market for medical interns and residents: a case study in game theory. *Journal of Political Economy*, 6(4):991–1016, 1984.
- [37] Roth A.E. On the allocation of residents to rural hospitals: a general property of two-sided matching markets. *Econometrica*, 54(2):425–427, 1986.
- [38] Roth A.E., and Peranson E. The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design. *American Economic Review*, 89:748–780, 1999.
- [39] Roth A.E., and Sotomayor M.A.O. Two-sided matching: a study in game-theoretic modeling and analysis. *Cambridge: Econometric Society monographs*, 1990.
- [40] Yokoi Y. A generalized polymatroid approach to stable matchings with lower quotas. *Mathematics of Operations Research*, 42(1):238–255, 2016.
- [41] Yokoi Y. Envy-Free Matchings with Lower Quotas. arXiv preprint, arXiv:1704.04888, 2017
- [42] Wu Q., and Roth A.E. The lattice of envy-free matchings. mimeo, 2016